

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:24 ; Search time 15.7143 Seconds  
(without alignments)  
97.882 Million cell updates/sec

Title: US-09-674-716B-3  
Perfect score: 81  
Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	112	1 KWS16	Ig kappa chain V r
2	81	100.0	113	1 KWS51	Ig kappa chain V r
3	81	100.0	120	1 KWS567	Ig kappa chain pre
4	81	100.0	120	2 A29775	Ig kappa chain pre
5	81	100.0	120	2 B29775	Ig kappa chain pre
6	66	81.5	112	2 PL0273	Ig kappa chain V r
7	65	80.2	133	2 S40324	Ig kappa chain V r
8	64	79.0	83	2 S34095	Ig kappa chain V r
9	64	79.0	111	2 S20709	Ig kappa chain V r
10	64	79.0	120	2 S42267	Ig kappa chain V r
11	64	79.0	120	2 S42268	Ig kappa chain V r
12	64	79.0	133	1 K2HURP	Ig kappa chain pre
13	63	77.8	114	2 S49572	Ig kappa chain pre
14	61	75.3	101	2 PH1057	Ig light chain V-I
15	61	75.3	113	1 K2HURF	Ig kappa chain V-I
16	61	75.3	122	2 S40338	Ig kappa chain V r
17	60	74.1	112	2 PL0275	Ig kappa chain V r
18	60	74.1	133	2 S42611	Ig kappa chain V r
19	60	74.1	249	2 S41374	HUNVK protein Fv an
20	59	72.8	87	2 S34094	Ig kappa chain V r
21	59	72.8	91	2 S42186	Ig kappa chain V r
22	59	72.8	101	2 A33730	Ig kappa chain V r
23	59	72.8	103	2 PH1055	Ig light chain V r
24	59	72.8	103	2 PH1056	Ig light chain V r
25	59	72.8	112	2 A36259	Ig kappa chain V r
26	59	72.8	113	2 F30560	Ig kappa chain V r
27	59	72.8	132	2 C32513	Ig kappa chain pre
28	59	72.8	133	2 S23230	Ig kappa chain pre
29	59	72.8	140	2 S22658	Ig kappa chain pre

ALIGNMENTS

RESULT 1

KWS16  
Ig kappa chain V region (M167) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Nov-1980 #sequence\_revision 30-Nov-1980 #text\_change 16-Aug-1996  
C:Accession: A01908  
R:Rudikoff, S.; Potter, M.  
Biochemistry 17, 2703-2707, 1978  
A:Title: kappa-Chain variable region from M167, a phosphorylcholine binding myeloma prot  
A:Reference number: A01908; MUID:79000273; PMID:99160  
A:Accession: A01908  
A:Molecule type: protein  
A:Residues: 1-112 <RUD>  
C:Comment: This chain was isolated from a myeloma protein that binds phosphorylcholine.  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer  
F:16-95/Domain: immunoglobulin homology <IMM>  
F:23-93/Disulfide bonds: #status predicted

Query Match 100.0%; Score 81; DB 1; Length 112;  
Best Local Similarity 100.0%; Pred. No. 3.le-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
Db 24 RSSKSLLYKDGKTYLN 39

RESULT 2

KWS51  
Ig kappa chain V region (M511) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 21-Jan-2000  
C:Accession: A01910  
R:Appella, E.  
Mol. Immunol. 17, 711-718, 1980  
A:Title: Amino acid sequence of the light chain variable region of M511, a phosphorylch  
A:Reference number: A01910; MUID:81052016; PMID:6776396  
A:Accession: A01910  
A:Molecule type: protein  
A:Residues: 1-113 <APP>  
C:Comment: This chain was isolated from a myeloma protein that binds phosphorylcholine.  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer  
F:16-95/Domain: immunoglobulin homology <IMM>  
F:23-93/Disulfide bonds: #status predicted

Query Match 100.0%; Score 81; DB 1; Length 113;

```

Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
Db 24 RSSKSLLYKDGKTYLN 39

RESULT 3
KWS67
Ig kappa chain precursor V region (VK167) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 21-Jan-2000
C:Accession: A01909
R:Jouvin-Marche, E.; Rudikoff, S.
Cell 25, 47-58, 1981
A:Title: Somatic mutation of immunoglobulin light-chain variable-region genes.
A:Reference number: A01909; MUID:82002223; PMID:6791832
A:Accession: A01909
A:Molecule type: DNA
A:Residues: 1-120 <SEL>
A>Note: the sequence was determined from the germline gene
C:Genetics:
A:Introns: 17/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into larger disulfide bonds.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-120/Product: Ig kappa chain V region (VK167) #status predicted <MAT>
F:36-115/Domain: immunoglobulin homology <IMM>
F:43-113/Disulfide bonds: #status predicted

Query Match 100.0%; Score 81; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
Db 44 RSSKSLLYKDGKTYLN 59

RESULT 4
A29775
Ig kappa chain precursor V region (mouse 24.2) - shrew mouse
C:Species: Mus pahari
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jan-2000
C:Accession: A29775
R:Jouvin-Marche, E.; Rudikoff, S.
Immunogenetics 24, 191-201, 1986
A:Title: Evolution of a V-kappa gene family.
A:Reference number: A91751; MUID:87006895; PMID:3093373
A:Accession: A29775
A:Molecule type: DNA
A:Residues: 1-120 <JOU>
A:Cross-references: GB:M1552; NID:gl97468; PIDN:AAA39036.1; PID:gl97469
A>Note: this sequence was determined from the germline gene
C:Genetics:
A:Introns: 17/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-120/Product: Ig kappa chain V region 24.2 #status predicted <MAT>
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 81; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
Db 44 RSSKSLLYKDGKTYLN 59

```

```

RESULT 5
B29775
Ig kappa chain precursor V region (mouse 24.1) - shrew mouse
C:Species: Mus pahari
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jan-2000
C:Accession: B29775
R:Jouvin-Marche, E.; Rudikoff, S.
Immunogenetics 24, 191-201, 1986
A:Title: Evolution of a V-kappa gene family.
A:Reference number: A91751; MUID:87006895; PMID:3093373
A:Accession: B29775
A:Molecule type: DNA
A:Residues: 1-120 <JOU>
A:Cross-references: GB:M1553; NID:gl97470; PIDN:AAA39037.1; PID:gl97471
A>Note: this sequence was determined from the germline gene
C:Genetics:
A:Introns: 17/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-120/Product: Ig kappa chain V region 24.1 #status predicted <MAT>
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 81; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
Db 44 RSSKSLLYKDGKTYLN 59

RESULT 6
PL0273
Ig kappa chain V region (anti-DNA, D23VK) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C:Accession: PL0273
R:Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein, A.
J. Exp. Med. 171, 265-297, 1990
A:Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic hypermutation
A:Reference number: PL0231; MUID:90111618; PMID:2104919
A:Accession: PL0273
A:Molecule type: mRNA
A:Residues: 1-112 <SHL>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-23/Region: framework 1
F:16-95/Domain: immunoglobulin homology <IMM>
F:24-39/Region: complementarity-determining 1
F:40-54/Region: framework 2
F:55-61/Region: complementarity-determining 2
F:62-93/Region: framework 3
F:94-102/Region: complementarity-determining 3
F:103-112/Region: framework 4

Query Match 81.5%; Score 66; DB 2; Length 112;
Best Local Similarity 75.0%; Pred. No. 0.00097;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
Db 24 RSSKSLLYKDGKTYLN 39

RESULT 7
S40324
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C>Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S40324
R:Klein, R.; Jaenichen, R.; Zachau, H.G.

```

```
Eur. J. Immunol. 23, 3248-3271, 1993
A:Title: Expressed human immunoglobulin chi genes and their hypermutation.
A:Reference number: S40312; MUID:94080891; PMID:8258341
A:Accession: S40324
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-133 <KLE>
A:Cross-references: EMBL:X72434
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:33-112/Domain: immunoglobulin homology <IMM>

Query Match      80.2%; Score 65; DB 2; Length 133;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
   |||:|||||
Db 41 RSDQSLVYSDGNTYLN 56

RESULT 8
S34095
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S34095
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387; PMID:8436174
A:Accession: S34095
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <WAG>
A:Cross-references: EMBL:X67179
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-80/Domain: immunoglobulin homology <IMM>

Query Match      79.0%; Score 64; DB 2; Length 83;
Best Local Similarity 75.0%; Pred. No. 0.0015;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
   |||:|||||
Db 9 RSDQSLVYSDGNTYLN 24

RESULT 9
S20709
Ig kappa chain V region - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jan-2000
C:Accession: S20709
R:Brennan, D.M.; Hinds, M.G.; Welsh, J.H.; Tempest, P.R.; Harris, W.J.; Carr, F.J.; Osb
submitted to the EMBL Data Library, April 1992
A:Description: Binding specificity and variable region sequences of two monoclonal antib
A:Reference number: S20706
A:Accession: S20709
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-111 <BRE>
A:Cross-references: EMBL:Z11917; NID:952655; PIDN:CAA77975.1; PID:952656
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-95/Domain: immunoglobulin homology <IMM>

Query Match      79.0%; Score 64; DB 2; Length 111;
Best Local Similarity 75.0%; Pred. No. 0.0021;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
```

```
Db 24 KSSQSLVYSDGKTYLN 39
   |||:|||||
RESULT 10
S42267
Ig kappa chain V region (A1) - human
C:Species: Homo sapiens (man)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: S42267
R:Lautner-Rieske, A.; Huber, C.; Meindl, A.; Pargent, W.; Schaeble, K.F.; Thiebe, R.; Zc
Eur. J. Immunol. 22, 1023-1029, 1992
A:Title: The human immunoglobulin kappa locus. Characterization of the duplicated A regi
A:Reference number: A49043; MUID:92201291; PMID:1551402
A:Accession: S42267
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <LAU>
A:Cross-references: EMBL:X63402
C:Genetics:
A:Introns: 17/3
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match      79.0%; Score 64; DB 2; Length 120;
Best Local Similarity 75.0%; Pred. No. 0.0022;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
   |||:|||||
Db 44 RSSQSLVYSDGNTYLN 59

RESULT 11
S42268
Ig kappa chain V region (A17) - human
C:Species: Homo sapiens (man)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jan-2000
C:Accession: S42268
R:Lautner-Rieske, A.; Huber, C.; Meindl, A.; Pargent, W.; Schaeble, K.F.; Thiebe, R.; Zo
Eur. J. Immunol. 22, 1023-1029, 1992
A:Title: The human immunoglobulin kappa locus. Characterization of the duplicated A regi
A:Reference number: A49043; MUID:92201291; PMID:1551402
A:Accession: S42268
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <LAU>
A:Cross-references: EMBL:X63403
C:Genetics:
A:Introns: 17/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match      79.0%; Score 64; DB 2; Length 120;
Best Local Similarity 75.0%; Pred. No. 0.0022;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
   |||:|||||
Db 44 RSSQSLVYSDGNTYLN 59

RESULT 12
K2HURP
Ig kappa chain precursor V-II region (RPM1) - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 15-Aug-1997
C:Accession: A01890
R:Klobeck, H.G.; Meindl, A.; Combrato, G.; Solomon, A.; Zachau, H.G.
Nucleic Acids Res. 13, 6499-6513, 1985
A:Title: Human immunoglobulin kappa light chain genes of subgroups II and III.
A:Reference number: A93588; MUID:86041852; PMID:2997711
A:Accession: A01890
```

A:Molecule type: DNA  
A:Residues: 1-133 <KLO>  
A>Note: the sequence was determined from the differentiated gene  
C:Genetics:  
A:Gene: GDB:IGKV2  
A:Cross-references: GDB:136265  
A:Map position: 2p12-2p12  
A:Introns: 1/11  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-133/Product: Ig kappa chain V-II region (RPMI) #status predicted <MAT>  
F:21-43/Region: framework 1  
F:36-115/Domain: immunoglobulin homology <IMM>  
F:44-59/Region: complementarity-determining 1  
F:60-74/Region: complementarity-determining 2  
F:75-81/Region: complementarity-determining 2  
F:82-113/Region: framework 3  
F:114-122/Region: complementarity-determining 3  
F:123-133/Region: framework 4  
F:43-113/Disulfide bonds: #status predicted

Query Match 79.0%; Score 64; DB 1; Length 133;  
Best Local Similarity 75.0%; Pred. No. 0.0025;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
Db 44 RSSQSLVYDGTNTYLN 59

RESULT 13  
S49572  
Ig kappa chain precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 08-Mar-1995 #sequence\_revision 14-Jul-1995 #text\_change 21-Jan-2000  
C:Accession: S49572  
R:Giachino, C.; Padovan, E.; Lanzavecchia, A.  
A:Description: k+1+ dual receptor B cells are present in the human peripheral repertoire  
A:Reference number: S49571  
A:Accession: S49572  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-114 <GIA>  
A:Cross-references: EMBL:Z46626; NID:G575261; PIDN:CAAB6596.1; PID:G575262  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
F:16-95/Domain: immunoglobulin homology <IMM>

Query Match 77.8%; Score 63; DB 2; Length 114;  
Best Local Similarity 75.0%; Pred. No. 0.0031;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
Db 24 RSSQSLVYDGTNTYLN 39

RESULT 14  
PH1057  
Ig light chain V region (clone 17a.145) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C:Accession: PH1057  
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A:Title: Both Igm and Igg anti-DNA antibodies are the products of clonally selective B c  
A:Reference number: PH0971; MUID:92381444; PMID:1512540  
A:Accession: PH1057  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA

A:Residues: 1-101 <TIL>  
A:Experimental source: B cell, strain [NZB x NZW]F1  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:16-94/Domain: immunoglobulin homology <IMM>

Query Match 75.3%; Score 61; DB 2; Length 101;  
Best Local Similarity 80.0%; Pred. No. 0.0059;  
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 15  
Db 24 RSSQSLVYDGTNTYLN 38

RESULT 15  
K2HUF8  
Ig kappa chain V-II region (Fr) - human (tentative sequence)  
C:Species: Homo sapiens (man)  
C:Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 31-Mar-2000  
C:Accession: A01886  
R:Riesen, W.F.; Jaton, J.C.  
Biochemistry 15, 3829-3833, 1976  
A:Title: Variable region sequence of the light chain from a Waldenstrom's IgM with spec  
A:Reference number: A01886; MUID:76253627; PMID:821524  
A:Accession: A01886  
A:Molecule type: protein  
A:Residues: 1-113 <RIE>  
C:Comment: This chain was isolated from a Waldenstrom's macroglobulin that binds phospho  
C:Genetics:  
A:Gene: GDB:IGKV2  
A:Cross-references: GDB:136265  
A:Map position: 2p12-2p12  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer  
F:16-95/Domain: immunoglobulin homology <IMM>  
F:23-93/Disulfide bonds: #status predicted

Query Match 75.3%; Score 61; DB 1; Length 113;  
Best Local Similarity 62.5%; Pred. No. 0.0067;  
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
Db 24 RSSQSLVYDGTNTYLN 39

Search completed: April 22, 2003, 12:54:45  
Job time: 16.7143 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:46:53 ; Search time 8 Seconds  
(without alignments)  
82.953 Million cell updates/sec

Title: US-09-674-716B-3

Perfect score: 81

Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	81	100.0	112	1 KV2A_MOUSE	P01626 mus musculus
2	81	100.0	113	1 KV2C_MOUSE	P01628 mus musculus
3	81	100.0	120	1 KV2B_MOUSE	P01627 mus musculus
4	64	79.0	133	1 KV2F_HUMAN	P06310 homo sapien
5	61	75.3	113	1 KV2B_HUMAN	P01615 homo sapien
6	55	67.9	113	1 KV2F_MOUSE	P01630 mus musculus
7	54	66.7	113	1 KV2G_MOUSE	P01631 mus musculus
8	53	65.4	113	1 KV2D_HUMAN	P01617 homo sapien
9	52	64.2	113	1 KV2E_MOUSE	P03376 mus musculus
10	46.5	57.4	115	1 KV2A_HUMAN	P01614 homo sapien
11	45	55.6	117	1 KV2E_HUMAN	P06309 homo sapien
12	43	53.1	437	1 ERR1_YEAST	P42222 saccharomyc
13	42	51.9	343	1 Y535_METJA	Q57955 methanococc
14	41	50.6	112	1 KV2D_MOUSE	P01629 mus musculus
15	41	50.6	891	1 ACON_LEGPN	P37032 legionella
16	40	49.4	87	1 R31B_PALSO	Q8Y062 ralteonias
17	40	49.4	478	1 AMY_STRHY	P08486 streptomyce
18	40	49.4	481	1 GPGD_DROME	P41572 drosophila
19	40	49.4	713	1 ETF2_YABAM	Q9qb84 yaba monkey
20	39.5	48.8	134	1 KV4C_HUMAN	P06314 homo sapien
21	39	48.1	213	1 ZN80_CERAE	P51502 cercopithe
22	39	48.1	293	1 ZN80_MACMU	P51505 macaca mula
23	39	48.1	473	1 SNACB_BACSU	P05655 bacillus su
24	39	48.1	478	1 NIFE_METH	O27607 methanobact
25	39	48.1	561	1 YBJL_ECOLI	P75812 escherichia
26	39	48.1	561	1 YBJL_SALTI	Q8z851 salmonella
27	39	48.1	561	1 YBJL_SALTI	Q8zqk4 salmonella
28	39	48.1	562	1 YD26_YERPE	Q8zgh1 yersinia pe
29	39	48.1	580	1 GGT_ECOLI	P18956 escherichia
30	38	46.9	135	1 Y545_RICPR	Q9zd01 rickettsia
31	38	46.9	213	1 VTXB_BACSU	P06568 bacillus su
32	38	46.9	216	1 YDCM_METS1	P43368 methylophil
33	38	46.9	262	1 LPXA_SALTI	Q8z9a2 salmonella

34	38	46.9	262	1 LPXA_SALTI	P32200 salmonella
35	38	46.9	262	1 LPXA_YERPE	Q8zhs6 yersinia pe
36	38	46.9	273	1 ZN80_GORGO	P51503 gorilla gor
37	38	46.9	273	1 ZN80_HUMAN	P51504 homo sapien
38	38	46.9	284	1 F100_LEIMA	P22045 leishmania
39	38	46.9	314	1 TOPI_VACCV	P08585 vaccinia vi
40	38	46.9	314	1 TOPI_VARV	P32989 variola vir
41	38	46.9	474	1 UCR2_EUGGR	P43265 euglena gra
42	38	46.9	506	1 TRPE_YEAST	P00899 saccharomyc
43	38	46.9	711	1 ETF2_MXVL	Q9g8k4 myxoma viru
44	38	46.9	711	1 ETF2_SFVKA	Q9g8y2 shope fibro
45	38	46.9	714	1 GP85_TRYCR	Q03877 trypanosoma

## ALIGNMENTS

RESULT 1  
KV2A\_MOUSE  
ID KV2A\_MOUSE STANDARD; PRT; 112 AA.  
AC P01626;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-II region MOPC 167.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=79000273; PubMed=99160;  
RA Rudikoff S, Potter M.;  
RT "kappa Chain variable region from M167, a phosphorylcholine binding  
myeloma protein.";  
RL Biochemistry 17:2703-2707(1978).  
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
BINDS PHOSPHORYLCHOLINE. THE SEQUENCE OF THE V REGION OF THE HEAVY  
CHAIN HAS ALSO BEEN DETERMINED.  
PIR: A01908; KWS16.  
DR HSSP; P80362; IWTIL.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig\_1.  
DR SMART; SM00406; IGV\_1.  
KW Immunoglobulin V region.  
FT DOMAIN 1 23  
FT DOMAIN 24 39  
FT DOMAIN 40 54  
FT DOMAIN 55 61  
FT DOMAIN 62 93  
FT DOMAIN 94 102  
FT DOMAIN 103 112  
FT DISULFID 23 93  
FT NON\_TER 112 112  
SQ SEQUENCE 112 AA; 12349 MW; A58EDFD6404B9726 CRC64;  
Query Match 100.0%; Score 81; DB 1; Length 112;  
Best Local Similarity 100.0%; Pred. NO. 2.7e-07;  
Matches 16; Conservative 0; Mismatches 0; Gaps 0;  
Indels 0;  
QY 1 RSSKSLLYKDGKTYLN 16  
DB 24 RSSKSLLYKDGKTYLN 39  
RESULT 2  
KV2C\_MOUSE  
ID KV2C\_MOUSE STANDARD; PRT; 113 AA.  
AC P01628;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

DE Ig kappa chain V-II region MOPC 511.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=81052016; PubMed=6776396;
RA Appella E.;
RT "Amino acid sequence of the light chain variable region of M511, a
RL phosphorolcholine-binding murine myeloma protein.";
RL Mol. Immunol. 17:711-718 (1980).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS PHOSPHORYLCHOLINE.
DR PIR; A01910; KVM551.
DR HSSP; P80362; IWL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT CHAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DISULFID 23 93
FT NON TER 113 113
SQ SEQUENCE 113 AA; 12496 MW; EPH0DC4DA2B03450 CRC64;

Query Match 100.0%; Score 81; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSSKSLLYKDGKTYLN 16
Db 24 RSSKSLLYKDGKTYLN 39

RESULT 3
KV2B MOUSE STANDARD; PRT; 120 AA.
AC P01627;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region VKAPPAL67 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82002223; PubMed=6791832;
RA Selsing E., Storb U.;
RT "Somatic mutation of immunoglobulin light-chain variable-region
RL genes.";
RL Cell 25:47-58 (1981).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00562; AAA39032.1; -.
DR EMBL; K02415; AAA39051.1; -.
DR PIR; A01909; KVM567.
DR HSSP; P80362; IWL.

```

```

DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 120
FT DOMAIN 21 43
FT DOMAIN 44 59
FT DOMAIN 60 74
FT DOMAIN 75 81
FT DOMAIN 82 113
FT DOMAIN 114 120
FT DISULFID 43 113
SQ SEQUENCE 120 AA; 13280 MW; 63BB571F0E4DE3E8 CRC64;

Query Match 100.0%; Score 81; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSSKSLLYKDGKTYLN 16
Db 44 RSSKSLLYKDGKTYLN 59

RESULT 4
KV2F HUMAN STANDARD; PRT; 133 AA.
AC P06310;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region RPMI 6410 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86041852; PubMed=2997711;
RA Klobbeck H.G., Weindl A., Combratio G., Solomon A., Zachau H.G.;
RT "Human immunoglobulin kappa light chain genes of subgroups II and
RL III.";
RL Nucleic Acids Res. 13:6499-6513 (1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z00020; CAA77315.1; -.
DR PIR; A01890; KZHURP.
DR HSSP; P80362; IWL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 133
FT DOMAIN 21 43
FT DOMAIN 44 59
FT DOMAIN 60 74
FT DOMAIN 75 81
FT DOMAIN 82 113
FT DOMAIN 114 122
FT DOMAIN 123 132
FT DISULFID 43 113
FT NON TER 133 133
SQ SEQUENCE 133 AA; 14707 MW; 513CCAF3673009EE CRC64;

```

```

Query Match      79.0%; Score 64; DB 1; Length 133;
Best Local Similarity 75.0%; Pred. No. 0.00032;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSSKSLLYKDGKTYLN 16
Db 44 RSSQSLVYSGNTYLN 59

RESULT 5
KV2B HUMAN          STANDARD;          PRT; 113 AA.
AC P01615;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region FR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE
RX MEDLINE=76253627; PubMed=821524;
RA Riessen W.F., Jaton J.-C.;
RT "Variable region sequence of the light chain from a Waldenstroms IgM
  with specificity for phosphorylcholine.";
RL Biochemistry 15:3829-3833(1976).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S
  MACROGLOBULIN THAT BINDS PHOSPHORYLCHOLINE.
DR PIR; A01886; K2HUPR.
DR HSSP; P01607; 1REL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DISULFID 23 93
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12660 MW; 0C0DA39E46DB96BE CRC64;

Query Match      75.3%; Score 61; DB 1; Length 113;
Best Local Similarity 62.5%; Pred. No. 0.00089;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSSKSLLYKDGKTYLN 16
Db 24 RSSQSLVYREGTYLB 39

RESULT 6
KV2F MOUSE          STANDARD;          PRT; 113 AA.
AC P01630;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 7S34.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE
RX MEDLINE=83256427; PubMed=6409088;

Query Match      75.3%; Score 61; DB 1; Length 113;
Best Local Similarity 62.5%; Pred. No. 0.00089;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSSKSLLYKDGKTYLN 16
Db 24 RSSQSLVYREGTYLB 39

RESULT 6
KV2F MOUSE          STANDARD;          PRT; 113 AA.
AC P01630;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 7S34.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE
RX MEDLINE=83256427; PubMed=6409088;

```

```

RA Chang J.-Y., Herbst H., Aebersold R., Braun D.G.;
RT "A new isotype sequence (V kappa 27) of the variable region of kappa-
  light chains from a mouse hybridoma-derived anti-(streptococcal group
  A polysaccharide) antibody containing an additional cysteine residue.
  Application of the dimethylaminoazobenzene isothiocyanate technique
  for the isolation of peptides.";
RL Biochem. J. 211:173-180(1983).
CC -1- MISCELLANEOUS: THIS CHAIN IS FROM A HYBRIDOMA-DERIVED MONOCLONAL
  ANTIBODY AGAINST THE STREPTOCOCCAL GROUP A POLYSACCHARIDE.
CC PIR; A01913; KWS57S.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Monoclonal antibody; Hybridoma.
FT DOMAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DISULFID 23 93
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12496 MW; 42C019D10ADA3C91 CRC64;

Query Match      67.9%; Score 55; DB 1; Length 113;
Best Local Similarity 73.3%; Pred. No. 0.01;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSSKSLLYKDGKTYL 15
Db 24 RSSKSLVHSGNTYL 38

RESULT 7
KV2G MOUSE          STANDARD;          PRT; 113 AA.
ID KV2G MOUSE
AC P01631;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 26-10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE
RX STRAIN=A/J;
RX MEDLINE=83178921; PubMed=6404298;
RA Novotny J., Margolies M.N.;
RT "Amino acid sequence of the light chain variable region from a mouse
  anti-digoxin hybridoma antibody.";
RL Biochemistry 22:1153-1158(1983).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG2A HYBRIDOMA
  PROTEIN THAT BINDS DIGOXIN.
CC PIR; A01914; KWS26.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Monoclonal antibody; Hybridoma.
FT DOMAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DISULFID 23 93
SQ SEQUENCE 113 AA; 12496 MW; 42C019D10ADA3C91 CRC64;

```

```

FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12273 MW; F9F39CE949A84C2A CRC64;

Query Match 66.7%; Score 54; DB 1; Length 113;
Best Local Similarity 62.5%; Pred. No. 0.015;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSKSLLYKDGKTYLN 16
   |||:::|::|::|
Db 24 RSSQSLVHSGNGTYLN 39

RESULT 8
KV2D_HUMAN
ID KV2D_HUMAN STANDARD; PRT; 113 AA.
AC P01617;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region TEW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (BENCE-JONES PROTEIN TEW).
RX MEDLINE=74148480; PubMed=4596149;
RA Putnam F.W., Whitley E.J. Jr., Paul C., Davidson J.N.;
RT "Amino acid sequence of a kappa Bence Jones protein from a case of
RL primary amyloidosis."
RL Biochemistry 12:3763-3780 (1973).
RN [2]
RP SEQUENCE OF 1-27 (AMYLOID PROTEIN TEW).
RX MEDLINE=73166638; PubMed=4700495;
RA Terry W.D., Page D.L., Kimura S., Isobe T., Osserman E.F.,
RT Glenner G.G.;
RT "Structural identity of Bence Jones and amyloid fibril proteins in a
RL patient with plasma cell dyscrasia and amyloidosis."
RL J. Clin. Invest. 52:1276-1281 (1973).
CC -1- MISCELLANEOUS: THE MAJOR AMYLOID PROTEIN APPEARS TO BE IDENTICAL
CC WITH THE BENCE JONES PROTEIN ISOLATED FROM THE SAME PATIENT.
CC -1- MISCELLANEOUS: THIS PROTEIN WAS ISOLATED FROM THE URINE OF A
CC PATIENT WITH PLASMA CELL DYSKRASIA AND AMYLOIDOSIS.
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)
CC MARKER.
DR PIR; A01888; K2HUTW.
DR HSSP; P01607; 1REI.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Bence-Jones protein; Amyloid.
FT DOMAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DOMAIN 113 113
FT DISULFID 23 93
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12316 MW; 0C3C38F81F1843CA CRC64;

Query Match 65.4%; Score 53; DB 1; Length 113;
Best Local Similarity 68.8%; Pred. No. 0.023;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RSKSLLYKDGKTYLN 16
   |||:::|::|::|
Db 24 RSSQSLVHSGNGTYLN 39

RESULT 9
KV2D_HUMAN
ID KV2D_HUMAN STANDARD; PRT; 113 AA.
AC P01617;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region TEW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=68242259; PubMed=5586923;
RA Hilschmann N.;
RT "The complete amino acid sequence of Bence Jones protein Cum (kappa-
RT type).";
RL Hoppe-Seyley's Z. Physiol. Chem. 348:1718-1722 (1967).
RN [2]
RP REVISIONS TO 50; 52; 96 AND 97.
RX MEDLINE=70063440; PubMed=4188189;
RA Hilschmann N.;
RT "Molecular basis of antibody formation.";
RL Naturwissenschaften 56:195-205 (1969).

```

```

KV2E_MOUSE
ID KV2E_MOUSE STANDARD; PRT; 113 AA.
AC P03976;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 17S29.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX TISSUE=Hybridoma;
RX MEDLINE=85128968; PubMed=6441768;
RA Abersold R., Herbst H., Grutter T., Chang J.Y., Braun D.G.;
RT "Murine V kappa 25 and V kappa 27 amino-acid sequences of C57Bl/6
RT origin: monoclonal antibodies 17S29.1 and 22S25.1 specific for the
RT group A-streptococcal polysaccharide."
RL Hoppe-Seyley's Z. Physiol. Chem. 365:1375-1381 (1984).
CC -1- MISCELLANEOUS: ANTI-STREPTOCOCCAL GROUP A CARBOHYDRATE ANTIBODY.
DR PIR; A01912; KWSI7.
DR HSSP; P01607; 1REI.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Hybridoma.
FT DOMAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DISULFID 23 93
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12390 MW; 4E93797046F8DB33 CRC64;

Query Match 64.2%; Score 52; DB 1; Length 113;
Best Local Similarity 73.3%; Pred. No. 0.034;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSKSLLYKDGKTYLN 15
   |||:::|::|::|
Db 24 RSSKSLVHSGNGITYLN 38

RESULT 10
KV2A_HUMAN
ID KV2A_HUMAN STANDARD; PRT; 115 AA.
AC P01614;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region Cum.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=68242259; PubMed=5586923;
RA Hilschmann N.;
RT "The complete amino acid sequence of Bence Jones protein Cum (kappa-
RT type).";
RL Hoppe-Seyley's Z. Physiol. Chem. 348:1718-1722 (1967).
RN [2]
RP REVISIONS TO 50; 52; 96 AND 97.
RX MEDLINE=70063440; PubMed=4188189;
RA Hilschmann N.;
RT "Molecular basis of antibody formation.";
RL Naturwissenschaften 56:195-205 (1969).

```



```

RESULT 13
Y535 METJA STANDARD; PRT; 343 AA.
AC Q57955;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0535.
GN MJ0535
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Klerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.P., Fuhmann J.L., Nguyen D.,
RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC ENBL; U67502; AAB98526.1; -.
DR TIGR; MJ0535; -.
DR InterPro; IPR000286; His deacetylase.
DR Pfam; PF00850; Hist deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
KW Hypothetical protein; Hydrolase; Complete proteome.
SQ SEQUENCE 343 AA; 38174 MW; 8848EDB5757FDC233 CRC64;

Query Match 51.9%; Score 42; DB 1; Length 343;
Best Local Similarity 69.2%; Pred. No. 6.5;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 SKSLLYKDGKTYL 15
| | | | |
DB 73 SKSNFYVDGTYL 85

RESULT 14
KV2D MOUSE
ID KV2D MOUSE STANDARD; PRT; 112 AA.
AC P01629;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 2S1.3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE.
RP MEDLINE=83055101; PubMed=7141411;
RA Herlitz H., Chang J.Y., Abersold R., Braun D.G.;
RT "Murine VK25 isotype sequence: monoclonal antibody 2S1.3 specific for

```

```

RT the group A streptococcal polysaccharide.";
RL Hoppe-Sevler's Z. Physiol. Chem. 363:1069-1076(1982).
CC -!- MISCELLANEOUS: THIS CHAIN IS FROM A HYBRIDOMA-DERIVED MONOCLONAL
CC PIR; A01911; KVMSS1.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Monoclonal antibody.
FT DOMAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DISULFID 23 93
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12221 MW; BD5EF5E6D789FBEC CRC64;

Query Match 50.6%; Score 41; DB 1; Length 112;
Best Local Similarity 66.7%; Pred. No. 2.8;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 RSSKSLYKDGKTYL 15
| | | | |
DB 24 RSSKSLYKDGKTYL 38

RESULT 15
ACON LEGN
ID ACON LEGN STANDARD; PRT; 891 AA.
AC P37032;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aconitate hydratase (EC 4.2.1.3) (Citrate hydro-lyase) (Aconitase)
DE (Major iron-containing protein) (MICP) (IP210).
DE ACN.
OS Legionella pneumophila.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=446;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Philadelphia 1;
RC MEDLINE=93374864; PubMed=8366052;
RX Mengaud J.M., Horwitz M.A.;
RT "The major iron-containing protein of Legionella pneumophila is an
RT aconitase homologous with the human iron-responsive element-binding
RT protein."
RL J. Bacteriol. 175:5666-5676(1993). Citrate = cis-aconitate + H(2)O.
CC -!- CATALYTIC ACTIVITY: Citrate = cis-aconitate + H(2)O.
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- MISCELLANEOUS: ACONITASE HAS AN ACTIVE (4FE-4S) AND AN INACTIVE
CC (3FE-4S) FORMS. THE ACTIVE (4FE-4S) CLUSTER IS PART OF THE
CC CATALYTIC SITE THAT INTERCONVERTS CITRATE, CIS-ACONITASE, AND
CC ISOCITRATE.
CC -!- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC ENBL; L22081; AAA25295.1; -.
DR PIR; B48642; B48642.

```

```

DR HSSP; P16276; 1B0J.
DR InterPro; IPR000573; Aconitase_C.
DR InterPro; IPR001030; Aconitase_N.
DR Pfam; PF00330; aconitase; 1.
DR Pfam; PF00694; Aconitase_C; 1.
DR PRINTS; PR00415; ACONITASE.
DR ProDom; PD000511; Aconitase_N; 1.
DR PROSITE; PS00450; ACONITASE_1; 1.
DR PROSITE; PS01244; ACONITASE_2; 1.
DR Lyase; Tricarboxylic acid cycle; Iron-sulfur; 4Fe-4S.
FT METAL 435 435 IRON (IRON-SULFUR CLUSTER)
      (BY SIMILARITY).
FT METAL 501 501 IRON (IRON-SULFUR CLUSTER)
      (BY SIMILARITY).
FT METAL 504 504 IRON (IRON-SULFUR CLUSTER)
      (BY SIMILARITY).
SQ SEQUENCE 891 AA; 98209 MW; F81FC99520859B48 CRC64;

Query Match          50.6%; Score 41; Length 891;
Best Local Similarity 69.2%; Pred. No. 28;
Matches 9; Conservative 1; Mismatchces 3; Indels

Qy 2 SSKSLLYDGTKY 14
   |::|::|::|::|
Db 9 STKSQLTVDGKTY 21

```

Search completed: April 22, 2003, 12:51:45  
Job time : 9 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:47:43 ; Search time 32 Seconds  
(without alignments)  
103.024 Million cell updates/sec

Title: US-09-674-716B-3

Perfect score: 81  
Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	75.3	239	4 Q8TCD0	Q8TCD0 homo sapien
2	57	70.4	114	4 Q9UL80	Q9UL80 homo sapien
3	50	61.7	758	10 Q9C9E8	Q9C9E8 arabidopsis
4	49	60.5	238	11 Q8VC16	Q8VC16 mus musculus
5	48	59.3	104	11 Q9JL82	Q9JL82 mus musculus
6	48	59.3	239	11 Q8VC55	Q8VC55 mus musculus
7	48	59.3	1561	17 Q26451	Q26451 methanobact
8	46	56.8	238	11 Q9M37	Q9M37 mus musculus
9	46	56.8	506	16 Q9L035	Q9L035 streptomyc
10	46	56.8	573	2 P97179	P97179 streptomyc
11	46	56.8	574	2 Q86876	Q86876 streptomyc
12	44	54.3	419	16 Q9CG26	Q9CG26 lactococcus
13	43	53.1	437	3 Q12007	Q12007 saccharomyc
14	43	53.1	481	3 Q9P8D3	Q9P8D3 sordaria ma
15	43	53.1	487	3 Q8X096	Q8X096 neurospora
16	43	53.1	489	3 Q9P8F3	Q9P8F3 gibberella

17	53.1	646	5	O96602	O96602 trypanosoma
18	53.1	822	16	O56939	O56939 versinia pe
19	51.9	239	16	Q8CMU7	Q8CMU7 pasteurella
20	51.9	245	16	O67538	O67538 aquifex ao
21	51.9	434	5	Q23217	Q23217 caenorhabdi
22	51.9	755	17	Q97VH4	Q97VH4 sulfolobus
23	51.9	786	5	O77209	O77209 trypanosoma
24	50.6	418	16	Q9XDT5	Q9XDT5 clostridium
25	50.6	555	16	Q8RIQ5	Q8RIQ5 fusobacteri
26	50.6	635	2	Q45820	Q45820 clostridium
27	50.6	651	16	Q9WYL6	Q9WYL6 thermotoga
28	50.6	700	2	O24781	O24781 bacillus sp
29	50.0	241	12	Q9YMQ6	Q9YMQ6 coltivirus
30	50.0	249	12	Q9YMQ2	Q9YMQ2 coltivirus
31	50.0	249	12	Q9YMP2	Q9YMP2 coltivirus
32	49.4	87	16	Q8Y062	Q8Y062 ralstonia s
33	49.4	127	16	Q9KAX6	Q9KAX6 bacillus ha
34	49.4	233	17	O59233	O59233 pyrococcus
35	49.4	305	16	O50934	O50934 borrelia bu
36	49.4	364	3	P78905	P78905 schizosacch
37	49.4	459	2	Q59964	Q59964 streptomyc
38	49.4	484	16	Q9KEM8	Q9KEM8 bacillus ha
39	49.4	486	10	Q9MZX0	Q9MZX0 arabidopsis
40	49.4	489	3	O94582	O94582 schizosacch
41	49.4	521	16	Q9PQU6	Q9PQU6 ureaplasma
42	49.4	761	13	Q918H4	Q918H4 opsanus tau
43	49.4	1016	2	Q9LCJ7	Q9LCJ7 leuconostoc
44	49.4	1928	5	Q9U0H2	Q9U0H2 plasmodium
45	39.5	48.8	257	16	Q8U9V9 agrobacteri

## ALIGNMENTS

### RESULT 1

ID	Q8TCD0	PRELIMINARY;	PRT;	239 AA.
AC	Q8TCD0;			
DT	01-JUN-2002	(TREMBLrel. 21, Created)		
DT	01-JUN-2002	(TREMBLrel. 21, Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DE	Hypothetical 26.2 kDa protein.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISUS=LUIG;			
RA	Strausberg R.;			
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC022362; AAH2362.1; -.			
KW	Hypothetical protein.			
SQ	SEQUENCE 239 AA; 26234 MW; FACEDC3A3B03871D CRC64;			

Query Match 75.3%; Score 61; DB 4; Length 239;  
Best Local Similarity 68.8%; Pred. NO. 0.029;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

### Qy 1 RSSKSLLYKDGKTYLN 16

Db 44 RSTQSLVSDGNTYLN 59

### RESULT 2

ID	Q9UL80	PRELIMINARY;	PRT;	114 AA.
AC	Q9UL80;			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	Myosin-reactive immunoglobulin light chain variable region (Fragment).			





```

ID Q9L035 PRELIMINARY; PRT; 506 AA.
AC Q9L035;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Secreted alpha-amylase.
GN AMLB OR SCO7020 OR SC1HL0.09.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Batsman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL352956; CAB88153.1; -.
DR HSSP; P29957; 1AQM.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR Pfam; PF00686; CBM_20; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR ProDom; PD001568; CBD_4; 1.
SQ SEQUENCE 506 AA; 53868 MW; FCC92A3BED9D2DAE CRC64;

Query Match 56.8%; Score 46; DB 16; Length 506;
Best Local Similarity 57.1%; Pred. No. 22;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTY 14
| : : : : :
Db 306 RNGETLYKDGATY 319

RESULT 10
P97179 PRELIMINARY; PRT; 573 AA.
ID P97179;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Alpha-amylase.

```

```

GN AMLB.
OS Streptomyces lividans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1916;
RN [1]
RP SEQUENCE FROM N.A.
RA Yin X.H., Gagnat J., Gerbaud C., Guerin M., Virolle M.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z85949; CAB06622.1; -.
DR EMBL; Z86113; CAB06815.1; -.
DR HSSP; P29957; 1AQM.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR Pfam; PF00686; CBM_20; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR ProDom; PD001568; CBD_4; 1.
SQ SEQUENCE 573 AA; 61214 MW; E474019661C9D6A2 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 573;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTY 14
| : : : : :
Db 306 RNGETLYKDGATY 319

RESULT 11
O86876 PRELIMINARY; PRT; 574 AA.
ID O86876;
AC O86876;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Alpha-amylase.
GN AML.
OS Streptomyces lividans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1916;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TK21;
RA Isiegas C., Parro V., Mellado R.P.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y13601; CAA73926.1; -.
DR HSSP; P29957; 1AQM.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR Pfam; PF00686; CBM_20; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR ProDom; PD001568; CBD_4; 1.
SQ SEQUENCE 574 AA; 61758 MW; A040AA1092C22270 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 574;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTY 14
| : : : : :
Db 306 RNGETLYKDGATY 319

RESULT 12
Q9CG26 PRELIMINARY; PRT; 419 AA.
ID Q9CG26;
AC Q9CG26;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

DR 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Diaminopimelate decarboxylase (EC 4.1.1.20).  
GN LYSA OR L11284.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
RX MEDLINE=21235186; PubMed=11337471;  
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,  
RA Weissenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis ssp. lactis IL1403.";  
RL Genome Res. 11:731-753(2001).  
DR EMBL; AE006360; AAK05382.1; --  
DR InterPro; IPR00183; Decarboxylase2.  
DR Pfam; PF02784; Orn\_Arg\_dec\_N; 1.  
DR Pfam; PF02784; Orn\_DAP\_Arg\_dec; 1.  
DR PRINTS; PR01179; ODADCRXLASE.  
KW Complete proteome.  
SQ SEQUENCE 419 AA; 47279 MW; F8305C3E1FAD71E9 CRC64;  
  
Query Match 54.3%; Score 44; DB 16; Length 419;  
Best Local Similarity 66.7%; Pred. No. 38;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RSSKSLLYKDGKTYL 15  
|||: ||: |||: ||  
Db 384 RSSEVLQEDGKTRL 398  
  
RESULT 13  
Q12007 PRELIMINARY; PRT; 437 AA.  
AC Q12007;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-  
DE glycerate hydro-lyase).  
GN (ERR1 OR YOR393W) AND (ERR2 OR YPL281C).  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE OF 1-331 FROM N.A.  
RA Delius H., Hebling U., Hofmann B.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 176-437 FROM N.A.  
RA Wambutt R., Wedler H.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Delius H., Hebling U.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Duesterhoeft A., Floeth M., Fritz M., Hilbert H., Moestl D.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: 2-PHOSPHO-D-GLYCERATE = PHOSPHOENOLPYRUVATE +  
CC H(2O).  
CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING  
CC THE DIMER (BY SIMILARITY).  
CC -1- PATHWAY: GLYCOLYSIS.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.  
DR EMBL; 275301; CAA99725.1; --

DR EMBL; 273637; CAA98018.1; --  
DR EMBL; 275302; CAA99728.1; --  
DR HSSP; P00924; 4ENL.  
DR SGD; S0005920; ERR1.  
DR InterPro; IPR000941; Enolase.  
DR Pfam; PF00113; enolase; 1.  
DR PRINTS; PR00148; ENOLASE.  
DR ProDom; PD000902; Enolase; 1.  
DR TIGRFAMs; TIGR01060; eno; 1.  
DR PROSITE; PS00164; ENOLASE; 1.  
KW Glycolysis; Lyase; Magnesium.  
SQ SEQUENCE 437 AA; 47327 MW; FAF09C00BE0E711C CRC64;  
  
Query Match 53.1%; Score 43; DB 3; Length 437;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 2 SSKSLLYKDGKTYLN 16  
||: ||: |||: ||  
Db 248 SAPSVFYKDGKTYLN 262  
  
RESULT 14  
Q9P8D3 PRELIMINARY; PRT; 481 AA.  
AC Q9P8D3;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE ATP citrate lyase, subunit 2 (EC 4.1.3.8).  
GN ACL2.  
OS Sordaria macrospora.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Sordaria.  
OX NCBI\_TaxID=5147;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3346; TISSUE=MYCELIUM;  
RA Newrousian M., Kuck U., Loser K., Weltring K.M.;  
RT "The fungal ac1 and ac12 genes encode two polypeptides with homology  
RT to the N- and C-terminal parts of the animal ATP citrate lyase  
RT polypeptide.";  
RL Curr. Genet. 0:0-0(0).  
DR EMBL; AJ243817; CAB76164.1; --  
KW Lyase.  
SQ SEQUENCE 481 AA; 52240 MW; 44C93D3C4567EB28 CRC64;  
  
Query Match 53.1%; Score 43; DB 3; Length 481;  
Best Local Similarity 60.0%; Pred. No. 65;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 2 SSKSLLYKDGKTYLN 16  
||: ||: |||: ||  
Db 2 SAKSILEADGKAILN 16  
  
RESULT 15  
Q8X096 PRELIMINARY; PRT; 487 AA.  
AC Q8X096;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Probable ATP citrate lyase subunit 2.  
GN B14D6.320.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,

RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora Genome project;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL356173; CAB91741.2; -.  
 KW Lyase.  
 SQ SEQUENCE 487 AA; 52969 MW; 56BC196FCCE4CBA2 CRC64;  
  
 Query Match 53.1%; Score 43; DB 3; Length 487;  
 Best Local Similarity 60.0%; Pred. NO. 66;  
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 2 SSKSLLYKDGKTYLN 16  
 |:|:| | | | |  
 DB 2 SAKSILEADGKAILN 16  
  
 Search completed: April 22, 2003, 12:53:45  
 Job time : 35 secs

GenCore version 5.1.4.p5\_4578  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:45:19 ; Search time 40.5714 Seconds

(without alignments)  
52.550 Million cell updates/sec

Title: US-09-674-716B-3

Perfect score: 81

Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 101002.\*

- 1: /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID82/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	81	100.0	16	21 AAY32254
2	81	100.0	100	22 AAE06969
3	81	100.0	116	21 AAY32262
4	81	100.0	145	21 AAY32261
5	77	95.1	16	19 AAW39818
6	77	95.1	113	19 AAW39882
7	77	95.1	113	19 AAW39804
8	76	93.8	12	AAR12232
9	76	93.8	132	12 AAR12354
10	75	92.6	16	19 AAW39815

11	75	92.6	16	19 AAW39824	Light chain CDR1 o
12	75	92.6	113	19 AAW39886	Light chain of the
13	75	92.6	113	19 AAW39801	Variable domain of
14	75	92.6	113	19 AAW39802	Variable domain of
15	75	92.6	113	19 AAW39803	Variable domain of
16	69	85.2	16	19 AAW39875	Light chain CDR1 o
17	69	85.2	16	19 AAW39839	Light chain CDR1 o
18	69	85.2	113	19 AAW39897	Light chain of the
19	69	85.2	113	19 AAW39805	Variable domain of
20	67	82.7	20	21 AAY70302	Murine anti-PAB-42
21	67	82.7	122	21 AAY70790	Murine anti-PAB-42
22	65	80.2	112	22 AAE06992	Human kappa light
23	64	79.0	93	21 AAY56670	Partial peptide fr
24	64	79.0	100	22 AAE06960	Mouse germline kap
25	64	79.0	100	22 AAE06961	Mouse germline kap
26	64	79.0	112	22 AAE06993	Light chain of a h
27	64	79.0	112	22 AAE06993	Human kappa light
28	64	79.0	112	22 AAE07003	Human kappa light
29	64	79.0	113	15 AAR52057	Light chain variab
30	64	79.0	113	15 AAR52057	Mouse monoclonal a
31	64	79.0	113	21 AAB12170	Mouse HBV pre-S1 a
32	64	79.0	113	21 AAB12171	Humanised HBV pre-
33	64	79.0	113	21 AAB12173	Humanised HBV pre-
34	64	79.0	132	20 AAV42306	Human anti-ganglio
35	64	79.0	133	21 AAY56735	Amino acid sequenc
36	64	79.0	139	21 AAY33724	The kappa chain of
37	64	79.0	247	23 AAU90900	Insulin/insulin-lf
38	64	79.0	253	23 ABP45953	Human Blys binding
39	62.5	77.2	15	19 AAW39821	Light chain CDR1 o
40	61	75.3	16	23 AAU70328	Human kappa II lig
41	61	75.3	112	22 AAE07007	Human kappa light
42	60	74.1	16	20 AAV14404	Peptide CDR-L1 der
43	60	74.1	112	20 AAY24099	Figure 11 protein
44	60	74.1	112	20 AAE06999	Human kappa light
45	59	72.8	16	22 AAB86292	Murine derived ant

## ALIGNMENTS

### RESULT 1

AAV32254

ID AAY32254 standard; Peptide; 16 AA.

XX AAY32254;

XX 15-FEB-2000 (first entry)

DE Light chain CDR L1 of mouse anti-CD23 MAB C11.

XX CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse;  
 KW monoclonal antibody; chimeric antibody; humanised antibody;  
 KW complementarity determining region; CDR; autoimmune disease;  
 KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
 KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
 KW urticaria; nephrotic syndrome; glomerulonephritis;  
 KW Sjogren's syndrome; allergy; ulcerative colitis; Crohn's disease;  
 KW graft-versus-host disease; COPD; bronchitis; diabetes;  
 KW B-cell malignancy; therapy.

OS Mus musculus.

PN WO958679-A1.

XX 18-NOV-1999.

XX 07-MAY-1999; 99WO-GB01434.

XX 09-MAY-1998; 98GB-0009839.

XX (GLAXO ) GLAXO GROUP LTD.

PI Bonney J, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 DR WPI; 2000-053101/04.  
 DR N-PSDB; AA234739.  
 XX Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 PT  
 XX  
 XX Claim 1; Page 40; 81pp; English.  
 XX  
 CC This sequence represents complementarity determining region 1  
 CC (CDR L1) of the light chain of murine anti-CD23 (FCER11) monoclonal  
 CC antibody C11 (see also AAY32262). The invention provides altered  
 CC antibodies, such as chimeric or humanised antibodies, which comprise  
 CC sufficient of the amino acid sequences of C11 light and heavy chain  
 CC CDRs (see AAY32254-59) to render them capable of binding to the CD23  
 CC type II molecule expressed on haematopoietic cells. The antibodies  
 CC are used to block soluble CD23 formation for treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis,  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX  
 SQ Sequence 16 AA;  
 Query Match 100.0%; Score 81; DB 21; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 5e-07;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RSSKSLLYKDGKTYLN 16  
 DB 1 RSSKSLLYKDGKTYLN 16  
 RESULT 2  
 AA06969  
 ID AAE06969 standard; Protein; 100 AA.  
 AC AAE06969;  
 XX  
 XX 16-OCT-2001 (first entry)  
 DT  
 DE Mouse germline kappa light chain variable (VK) region, 167/24.  
 KW Mouse; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
 KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
 KW multiple sclerosis; atherosclerosis; atherosclerosis; restenosis; asthma;  
 KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
 KW AIDS; inflammatory glomerulopathy; vascular intervention;  
 KW neointimal hyperplasia; VK; kappa light chain variable region.  
 XX  
 OS Mus sp.  
 XX  
 XX WO200157226-A1.  
 FN  
 PD 09-AUG-2001.  
 XX  
 PF 02-FEB-2001; 2001WO-US03537.  
 XX  
 PR 03-FEB-2000; 2000US-0497625.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Larosa GS, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;

XX WPI; 2001-488888/53.  
 XX  
 DR Humanized immunoglobulin for treating a CC-chemokine receptor  
 PT 2-mediated disorder in a patient, comprises a binding specificity for  
 FT CCR2, and a non-human antigen binding region and human immunoglobulin  
 FT  
 XX  
 XX Disclosure; Page 151; 183pp; English.  
 PS  
 XX  
 CC The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherosclerosis and arteriosclerosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is mouse germline kappa light chain variable  
 CC (VK) region, 167/24.  
 XX  
 SQ Sequence 100 AA;  
 Query Match 100.0%; Score 81; DB 22; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RSSKSLLYKDGKTYLN 16  
 DB 24 RSSKSLLYKDGKTYLN 39  
 RESULT 3  
 AA32262  
 ID AAY32262 standard; Protein; 116 AA.  
 AC AAY32262;  
 XX  
 XX 15-FEB-2000 (first entry)  
 DT  
 DE Humanised anti-CD23 MAb C11 light chain variable region.  
 KW CD23; FCER11; IgE receptor; monoclonal antibody; C11; mouse; human;  
 KW monoclonal antibody; chimeric antibody; humanised antibody;  
 KW complementarity determining region; CDR; autoimmune disease;  
 KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
 KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
 KW urticaria; nephrotic syndrome; glomerulonephritis;  
 KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
 KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
 KW graft-versus-host disease; COPD; bronchitis; diabetes;  
 KW B-cell malignancy; therapy.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT 1..23  
 FT Region /note= "framework region 1"  
 FT 24..39



```

FT Region /note= "CDR 1"
FT 40..54
FT /note= "framework region 2"
FT 55..61
FT /note= "CDR 2"
FT 62..93
FT /note= "framework region 3"
FT 94..102
FT /note= "CDR 3"
FT 103..113
FT /note= "framework region 4"
FT 114..124
XX WO958679-A1.
XX 18-NOV-1999.
XX 07-MAY-1999; 99WO-GB01434.
XX 09-MAY-1998; 98GB-0009839.
XX (GLAX ) GLAXO GROUP LTD.
XX Bonneyfey JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;
XX WPI; 2000-053101/04.
XX N-PSDB; AA234747.
XX Cell receptor specific antibodies useful for treating e.g. arthritis,
XX diabetes, multiple sclerosis and psoriasis -
XX Claim 9; Fig 3; 81pp; English.
XX This sequence represents the light chain variable region (VL) of
XX humanised anti-CD23 (FCER1I) monoclonal antibody C11, composed of
XX a human framework (H5IGKVI) and the light chain complementarity
XX determining regions (see AAY32254-56) of murine antibody C11. The
XX DNA was constructed by splice overlap PCR. The invention provides
XX altered antibodies, such as chimeric or humanised antibodies, which
XX comprise sufficient of the amino acid sequences of the C11 light
XX and heavy chain complementarity determining regions to render them
XX capable of binding to the CD23 type II molecule expressed on
XX haematopoietic cells. The antibodies are used to block soluble
XX CD23 formation in human therapy, for the treatment of arthritis,
XX lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,
XX diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic
XX syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative
XX colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic
XX asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,
XX eczema, graft-versus-host disease, COPD, insulinitis, bronchitis
XX (particularly chronic bronchitis) or diabetes (particularly type 1
XX diabetes), and B-cell malignancies (claimed). They are also useful
XX for studying interactions between CD23 and various ligands and
XX determining the binding agents.
XX Sequence 116 AA;
XX Query Match 100.0%; Score 81; DB 21; Length 116;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-06;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RSSKSLLYKDGKTYLN 16
XX |||||
XX Db 24 RSSKSLLYKDGKTYLN 39
XX
XX RESULT 4
XX ID AAY32261
XX AC AAY32261;
XX XX 15-FEB-2000 (first entry)
XX

```

```

DE Mouse anti-CD23 Mab C11 light chain variable region.
XX
XX CD23; FCER1I; IGE receptor; monoclonal antibody; C11; mouse;
XX monoclonal antibody; chimeric antibody; humanised antibody;
XX complementarity determining region; CDR; autoimmune disease;
XX inflammation; arthritis; lupus erythematosus; multiple sclerosis;
XX Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;
XX urticaria; nephrotic syndrome; glomerulonephritis;
XX inflammatory bowel disease; ulcerative colitis; Crohn's disease;
XX Sjogren's syndrome; allergies; asthma; rhinitis; eczema; insulinitis;
XX graft-versus-host disease; COPD; bronchitis; diabetes;
XX B-cell malignancy; therapy.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX Region 55..70
XX /note= "CDR L1"
XX Region 83..92
XX /note= "CDR L2"
XX Region 125..134
XX /note= "CDR L3"
XX
XX WO958679-A1.
XX 18-NOV-1999.
XX 07-MAY-1999; 99WO-GB01434.
XX 09-MAY-1998; 98GB-0009839.
XX (GLAX ) GLAXO GROUP LTD.
XX Bonneyfey JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;
XX WPI; 2000-053101/04.
XX N-PSDB; AA234746.
XX Cell receptor specific antibodies useful for treating e.g. arthritis,
XX diabetes, multiple sclerosis and psoriasis -
XX Claim 8; Fig 2; 81pp; English.
XX This sequence represents the light chain variable region (VL) of
XX murine anti-CD23 (FCER1I) monoclonal antibody C11. The invention
XX provides altered antibodies, such as chimeric or humanised
XX antibodies (see AAY32262 and AAY32263), which comprise sufficient of
XX the amino acid sequences of the C11 light and heavy chain
XX complementarity determining regions (see AAY32254-59) to render them
XX capable of binding to the CD23 type II molecule expressed on
XX haematopoietic cells. The antibodies are used to block soluble
XX CD23 formation in human therapy, for the treatment of arthritis,
XX lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,
XX diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic
XX syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative
XX colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic
XX asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,
XX eczema, graft-versus-host disease, COPD, insulinitis, bronchitis
XX (particularly chronic bronchitis) or diabetes (particularly type 1
XX diabetes), and B-cell malignancies (claimed). They are also useful
XX for studying interactions between CD23 and various ligands and
XX determining the binding agents.
XX Sequence 145 AA;
XX Query Match 100.0%; Score 81; DB 21; Length 145;
XX Best Local Similarity 100.0%; Pred. No. 5.8e-06;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RSSKSLLYKDGKTYLN 16
XX |||||
XX Db 55 RSSKSLLYKDGKTYLN 70
XX

```

```

RESULT 5
AAW39818
ID AAW39818 standard; peptide; 16 AA.
XX AC AAW39818;
XX DT 16-JUN-1998 (first entry)
XX DE Light chain CDR1 of catalytic antibody 6A12.
XX KW Variable domain; lambda light chain; catalytic antibody; degradation;
XX KW cocaine; cocaine transition state analogue; TSA; benzoic acid;
XX KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;
XX KW overdose; addiction.
XX OS Mus sp.
XX PN WO9749800-A1.
XX PD 31-DEC-1997.
XX PF 25-JUN-1997; 97WO-US10965.
XX PR 25-JUN-1996; 96US-0672345.
XX PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX PI Landry DW;
XX DR WPI; 1998-077166/07.
XX PT New catalytic antibodies able to decompose cocaine, single-chain
XX PT analogues used to treat cocaine overdose and addiction, required
XX PT in far smaller doses than antibodies that antagonise cocaine by
XX PT simply binding
XX PS Claim 13; Page 81; 147pp; English.
XX CC AAW39818-20 represent the sequences of the light chain complementarity
XX CC determining regions (CDRs) of the catalytic antibody 6A12, which is able
XX CC to degrade cocaine. A series of cocaine transition state analogues
XX CC (TSAs) were prepared and used to immunise mice for production of
XX CC hybridomas. Catalytic antibodies were identified by their capacity to
XX CC release 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody was
XX CC identified using TSA1, which is an immunogenic conjugate of a phosphate
XX CC monoester transition state analogue. Antibody 6A12 has a per minute Kcat
XX CC of 0.072. The antibodies reduce the concentration of cocaine in a
XX CC subject, and are used particularly for the treatment of cocaine in a
XX CC are also used for treating addiction (by reducing the in vivo
XX CC concentration that can be achieved).
XX SQ Sequence 16 AA;
Query Match 95.1%; Score 77; DB 19; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.4e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYDKGKTYLN 16
Db |||||:|||||
1 RSSKSLLYEDGKTYLN 16

RESULT 6
AAW39882
ID AAW39882 standard; Protein; 113 AA.
XX AC AAW39882;
XX DT 16-JUN-1998 (first entry)
XX DE Light chain of the catalytic antibody 2A10.
XX KW Variable domain; lambda light chain; catalytic antibody; degradation;
XX KW cocaine; cocaine transition state analogue; TSA; benzoic acid;
XX KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;
XX KW overdose; addiction.
XX OS Mus sp.
XX PN WO9749800-A1.
XX PD 31-DEC-1997.
XX PF 25-JUN-1997; 97WO-US10965.
XX PR 25-JUN-1996; 96US-0672345.
XX PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX PI Landry DW;
XX DR WPI; 1998-077166/07.
XX PT New catalytic antibodies able to decompose cocaine, single-chain
XX PT analogues used to treat cocaine overdose and addiction, required
XX PT in far smaller doses than antibodies that antagonise cocaine by
XX PT simply binding
XX PS Disclosure; Fig 21; 147pp; English.
XX CC The present sequence represents the light chain of a catalytic antibody
XX CC which is capable of degrading cocaine. A series of cocaine transition
XX CC state analogues (TSAs) were prepared and used to immunise mice for
XX CC production of hybridomas. Catalytic antibodies were identified by their
XX CC capacity to release 3H-benzoic acid from 3H-phenyl cocaine. The
XX CC antibodies reduce the concentration of cocaine in a subject, and are used
XX CC particularly for the treatment of an overdose. They are also used for
XX CC treating addiction (by reducing the in vivo concentration that can be
XX CC achieved).
XX SQ Sequence 113 AA;
Query Match 95.1%; Score 77; DB 19; Length 113;
Best Local Similarity 93.8%; Pred. No. 2.1e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYDKGKTYLN 16
Db |||||:|||||
24 RSSKSLLYEDGKTYLN 39

RESULT 7
AAW39804
ID AAW39804 standard; protein; 113 AA.
XX AC AAW39804;
XX DT 16-JUN-1998 (first entry)
XX DE Variable domain of the Kappa light chain of catalytic antibody 2A10.
XX KW Variable domain; lambda light chain; catalytic antibody; degradation;
XX KW cocaine; cocaine transition state analogue; TSA; benzoic acid;
XX KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;
XX KW overdose; addiction.
XX OS Mus sp.
XX PN WO9749800-A1.
XX PD 31-DEC-1997.
XX PF 25-JUN-1997; 97WO-US10965.

```

PR 25-JUN-1996; 96US-0672345.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Landry DW;  
XX WPI; 1998-077166/07.  
XX  
XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding  
XX  
XX Claim 16; Pages 73-74; 147pp; English.  
XX  
XX AA139801-05 represent the amino acid sequences of the variable domain  
CC of the Kappa light chain of catalytic antibodies which are able to  
CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. The 2A10 antibody (AA139809  
CC represents the heavy chain) was identified using T9A1, which is an  
CC immunogenic conjugate of a phosphate monoester transition state  
CC analogue. Antibody 2A10 has a per minute Kcat of 0.011. The antibodies  
CC reduce the concentration of cocaine in a subject, and are used  
CC particularly for the treatment of an overdose. They are also used  
CC for treating addiction (by reducing the in vivo concentration that can  
CC be achieved).  
XX  
XX Sequence 113 AA;  
SQ  
Query Match 95.1%; Score 77; DB 19; Length 113;  
Best Local Similarity 93.8%; Pred. No. 2.1e-05; Indels 0; Gaps 0;  
Matches 15; Conservative 1; Mismatches 0;  
Qy 1 RSSKSLLYKDGKTYLN 16  
Db 24 RSSKSLLYEDGKTYLN 39  
|||||:|||||  
RESULT 8  
AA12232  
ID AAR12232 standard; Protein; 131 AA.  
XX  
XX AAR12232;  
XX  
XX 19-AUG-1991 (first entry)  
XX  
XX Mouse MAb 2E12 L chain V region.  
XX  
XX HIV-1; chimera.  
XX  
XX Mus sp.  
XX  
XX WO9107494-A.  
XX  
XX 30-MAY-1991.  
XX  
XX 13-NOV-1990; 90WO-US06627.  
XX  
XX 13-NOV-1989; 89US-0433703.  
XX  
XX (XOMA-) XOMA CORP.  
XX  
XX Better MD, Horwitz AH, Ghosh-Dastidar P, Robinson RR;  
XX WPI; 1991-178106/24.  
XX N-PSDB; AAQ12012.  
XX  
XX New chimeric mouse human antibodies - used in treatment, diagnosis  
PT and prophylaxis of HIV infections.  
XX  
XX Disclosure; Fig 1; 108pp; English.  
PS

XX The mouse VL gene product may be used to produce chimeric mouse-  
CC human Abs against HIV-1 comprising human Ig constant regions and  
CC murine variable regions. These novel sequence are useful in  
CC treatment, diagnosis and prophylaxis of HIV infections, and may be  
CC produced by a bacterial, yeast or mammalian expression system.  
XX  
XX Sequence 131 AA;  
SQ  
Query Match 93.8%; Score 76; DB 12; Length 131;  
Best Local Similarity 93.8%; Pred. No. 3.7e-05; Indels 0; Gaps 0;  
Matches 15; Conservative 1; Mismatches 0;  
Qy 1 RSSKSLLYKDGKTYLN 16  
Db 44 RSSKSLLYKDGKTYLS 59  
|||||:|||||  
RESULT 9  
AAR12354  
ID AAR12354 standard; Protein; 132 AA.  
XX  
XX AAR12354;  
XX  
XX 15-AUG-1991 (first entry)  
XX  
XX Light (kappa) chain variable region of murine 2E12  
DE immunoglobulin.  
XX  
XX Chimeric antibodies; immunoconjugates; HIV; AIDS.  
XX  
XX Mus musculus.  
XX  
XX WO9107493-A.  
XX  
XX 30-MAY-1991.  
XX  
XX 13-NOV-1990; 90WO-US06615.  
XX  
XX 13-NOV-1989; 89US-0433730.  
XX  
XX (XOMA-) XOMA CORP.  
XX  
XX (GREC ) GREEN CROSS CORP.  
XX  
XX Better MD, Horwitz AH, Ghosh-Dastidar P, Robinson RR;  
XX WPI; 1991-178044/24.  
XX N-PSDB; AAQ12056.  
XX  
XX New chimeric mouse-human antibodies - used to detect, kill and  
PT remove HIV-1 antigen from sample  
XX  
XX Disclosure; fig 1; 107pp; English.  
XX  
XX This is the light (kappa) - chain variable (V) region of a mouse  
CC monoclonal antibody (Mab), 2E12, and is specific for an HIV-1  
CC viral antigen. It is used in the construction of a chimeric Mab  
CC comprising heavy and light chains having murine V regions and human  
CC C regions. The chimeric MABs are more effective than murine MAB  
CC 2E12 since they have an increased compatibility in humans. The  
CC heavy and light chain V-regions are joined by manipulating their  
CC respective joining (J) regions, to generate restriction enzyme  
CC recognition sites. The chimeric MABs can be used as immuno-  
CC conjugates, in association with e.g. toxins for HIV treatment.  
CC They can also be used in diagnosis of HIV.  
CC See also AAQ12057-63.  
XX  
XX Sequence 132 AA;  
SQ  
Query Match 93.8%; Score 76; DB 12; Length 132;  
Best Local Similarity 93.8%; Pred. No. 3.7e-05; Indels 0; Gaps 0;  
Matches 15; Conservative 1; Mismatches 0;  
Qy 1 RSSKSLLYKDGKTYLN 16



PD 31-DEC-1997.  
 XX 25-JUN-1997; 97WO-US10965.  
 XX 25-JUN-1996; 96US-0672345.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Landry DW;  
 XX WPI; 1998-077166/07.  
 XX N-PSDB; AAV09793.  
 XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX Disclosure; Fig 19; 147pp; English.  
 XX The present sequence represents the light chain of a catalytic antibody  
 CC which is capable of degrading cocaine. A series of cocaine transition  
 CC state analogues (TSAs) were prepared and used to immunise mice for  
 CC production of hybridomas. Catalytic antibodies were identified by their  
 CC capacity to release 3H-benzoic acid from 3H-phenyl cocaine. The  
 CC antibodies reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used for  
 CC treating addiction (by reducing the in vivo concentration that can be  
 CC achieved).  
 XX Sequence 113 AA;  
 SQ  
 Query Match 92.6%; Score 75; DB 19; Length 113;  
 Best Local Similarity 87.5%; Pred. No. 4.6e-05;  
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RSSKSLLYKDGKTYLN 16  
 DB 24 RSSRSLLYRDGKTYLN 39  
 RESULT 13  
 AAW39801  
 ID AAW39801 standard; protein; 113 AA.  
 XX AC AAW39801;  
 XX 16-JUN-1998 (first entry)  
 XX Variable domain of the Kappa light chain of catalytic antibody 3B9.  
 XX Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX Mus sp.  
 OS WO9749800-A1.  
 PN 31-DEC-1997.  
 PD 25-JUN-1997; 97WO-US10965.  
 XX 25-JUN-1996; 96US-0672345.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Landry DW;  
 XX WPI; 1998-077166/07.  
 XX P-PSDB; AAV09791.  
 XX

PT New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX Claim 12; Pages 71-72; 147pp; English.  
 XX AAW39801-05 represent the amino acid sequences of the variable domain  
 CC of the Kappa light chain of catalytic antibodies which are able to  
 CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
 CC were prepared and used to immunise mice for production of hybridomas.  
 CC Catalytic antibodies were identified by their capacity to release  
 CC 3H-benzoic acid from 3H-phenyl cocaine. The 3B9 antibody (AAW39806  
 CC represents the heavy chain) was identified using TSA1, which is an  
 CC immunogenic conjugate of a phosphate monoester transition state  
 CC analogue. Antibody 3B9 has a per minute Kcat of 0.11. The antibodies  
 CC reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used  
 CC for treating addiction (by reducing the in vivo concentration that can  
 CC be achieved).  
 XX Sequence 113 AA;  
 SQ  
 Query Match 92.6%; Score 75; DB 19; Length 113;  
 Best Local Similarity 87.5%; Pred. No. 4.6e-05;  
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RSSKSLLYKDGKTYLN 16  
 DB 24 RSSRSLLYRDGKTYLN 39  
 RESULT 14  
 AAW39802  
 ID AAW39802 standard; protein; 113 AA.  
 XX AC AAW39802;  
 XX 16-JUN-1998 (first entry)  
 XX Variable domain of the Kappa light chain of catalytic antibody 6A12.  
 XX Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX Mus sp.  
 OS WO9749800-A1.  
 PN 31-DEC-1997.  
 PD 25-JUN-1997; 97WO-US10965.  
 XX 25-JUN-1996; 96US-0672345.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Landry DW;  
 XX WPI; 1998-077166/07.  
 XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX Claim 14; Page 72; 147pp; English.  
 XX AAW39801-05 represent the amino acid sequences of the variable domain  
 CC of the Kappa light chain of catalytic antibodies which are able to  
 CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
 CC

CC were prepared and used to immunise mice for production of hybridomas.  
 CC Catalytic antibodies were identified by their capacity to release  
 CC 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody (AAW39807  
 CC represents the heavy chain) was identified using TSA1, which is an  
 CC immunogenic conjugate of a phosphate monoester transition state  
 CC analogue. Antibody 6A12 has a per minute Kcat of 0.072. The antibodies  
 CC reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used  
 CC for treating addiction (by reducing the in vivo concentration that can  
 CC be achieved).

XX SQ Sequence 113 AA;

Query Match 92.6%; Score 75; DB 19; Length 113;  
 Best Local Similarity 87.5%; Pred. No. 4.6e-05;  
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
 |||:||||:|||||  
 DB 24 RSSRSLLYRDGKTYLN 39

## RESULT 15

AAW39803  
 ID AAW39803 standard; protein; 113 AA.

XX AC AAW39803;

XX DT 16-JUN-1998 (first entry)

XX DE Variable domain of the kappa light chain of catalytic antibody 12H1.

XX KW Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.

XX OS MUB sp.

XX PN WO9749800-A1.

XX PD 31-DEC-1997.

XX PF 25-JUN-1997; 97WO-US10965.

XX PR 25-JUN-1996; 96US-0672345.

XX PA (UYCO ) UNIV COLUMBIA NEW YORK.

XX PI Landry DW;

XX DR WPI; 1998-077166/07.

XX DR P-PSDB; AAW09802.

XX PT New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding

XX PS Claim 18; Page 73; 147pp; English.

XX CC AAW39801-05 represent the amino acid sequences of the variable domain  
 CC of the kappa light chain of catalytic antibodies which are able to  
 CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
 CC were prepared and used to immunise mice for production of hybridomas.  
 CC Catalytic antibodies were identified by their capacity to release  
 CC 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody (AAW39808  
 CC represents the heavy chain) was identified using TSA2, and has a per  
 CC minute Kcat of 0.016. The antibodies reduce the concentration of cocaine  
 CC in a subject, and are used particularly for the treatment of an  
 CC overdose. They are also used for treating addiction (by reducing the in  
 CC vivo concentration that can be achieved).

SQ Sequence 113 AA;

Query Match 92.6%; Score 75; DB 19; Length 113;  
 Best Local Similarity 87.5%; Pred. No. 4.6e-05;  
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16

|||:||||:|||||

DB 24 RSSRSLLYRDGKTYLN 39

Search completed: April 22, 2003, 12:51:11  
 Job time : 41.5714 secs

GenCore version 5.1.4.p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:53:59 / Search time 16.8571 Seconds  
(without alignments)  
76.055 Million cell updates/sec

Title: US-09-674-716B-3

Perfect score: 81

Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	81	100.0	100	10	US-09-840-459-32
2	65	80.2	112	10	US-09-840-459-55
3	64	79.0	93	10	US-09-305-243-61
4	64	79.0	100	9	US-10-194-975-75
5	64	79.0	100	9	US-10-194-975-76
6	64	79.0	100	10	US-09-840-459-23
7	64	79.0	100	10	US-09-840-459-24
8	64	79.0	112	10	US-09-840-459-56
9	64	79.0	112	10	US-09-840-459-66
10	64	79.0	132	12	US-10-006-773-11
11	64	79.0	222	9	US-09-479-614-26
12	64	79.0	242	9	US-09-479-614-20
13	64	79.0	253	9	US-09-880-748-1964
14	61	75.3	112	10	US-09-840-459-70
15	60	74.1	112	10	US-09-840-459-62
16	59	72.8	100	10	US-09-840-459-22
17	59	72.8	111	10	US-09-835-087-2
18	59	72.8	111	10	US-09-809-739-13
19	59	72.8	111	10	US-09-840-459-11

20	59	72.8	111	10	US-09-840-459-59	Sequence 59, Appl
21	59	72.8	112	9	US-10-032-482-8	Sequence 8, Appl
22	59	72.8	112	10	US-09-772-120-6	Sequence 6, Appl
23	59	72.8	112	10	US-09-840-459-54	Sequence 54, Appl
24	59	72.8	112	10	US-09-840-459-58	Sequence 58, Appl
25	59	72.8	113	10	US-09-840-459-63	Sequence 63, Appl
26	59	72.8	244	9	US-09-880-748-1991	Sequence 1991, Ap
27	59	72.8	535	9	US-09-968-851-38	Sequence 38, Appl
28	59	71.6	100	9	US-10-194-975-78	Sequence 78, Appl
29	58	71.6	112	10	US-09-840-459-60	Sequence 60, Appl
30	58	71.6	113	10	US-09-840-459-64	Sequence 64, Appl
31	58	71.6	113	10	US-09-840-459-65	Sequence 65, Appl
32	57	70.4	112	10	US-09-840-459-67	Sequence 67, Appl
33	56	69.1	112	10	US-09-840-459-69	Sequence 69, Appl
34	55	67.9	16	10	US-09-835-087-14	Sequence 14, Appl
35	55	67.9	100	10	US-09-840-459-21	Sequence 21, Appl
36	55	67.9	112	10	US-09-835-087-1	Sequence 1, Appl
37	55	67.9	112	10	US-09-835-087-3	Sequence 3, Appl
38	55	67.9	112	10	US-09-835-087-4	Sequence 4, Appl
39	55	67.9	112	10	US-09-835-087-5	Sequence 5, Appl
40	55	67.9	112	10	US-09-835-087-6	Sequence 6, Appl
41	55	67.9	112	10	US-09-835-087-7	Sequence 7, Appl
42	55	67.9	112	10	US-09-809-739-11	Sequence 11, Appl
43	55	67.9	112	10	US-09-809-739-14	Sequence 14, Appl
44	55	67.9	112	10	US-09-809-739-15	Sequence 15, Appl
45	55	67.9	112	10	US-09-809-739-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1  
US-09-840-459-32  
; Sequence 32, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: Larrosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran H.  
; APPLICANT: O'Brien, Siobhan H.  
; APPLICANT: O'Keefe, Theresa  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; \*TITLE OF INVENTION: METHODS OF USE THEREFOR  
; FILE REFERENCE: 1855.1052-012  
; CURRENT APPLICATION NUMBER: US/09/840,459  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 09/497,625  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 09/359,193  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 09/121,781  
; PRIOR FILING DATE: 1998-07-23  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 32  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-840-459-32

Query Match 100.0%; Score 81; DB 10; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSSKSLLYKDGKTYLN 16

Db 24 RSSKSLLYKDGKTYLN 39

RESULT 2





Thu Apr 24 09:13:30 2003

us-09-674-716b-3.open.rapb

;; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
;; FILE REFERENCE: 1855.1052-012  
;; CURRENT APPLICATION NUMBER: US/09/840,459

;; PRIORITY FILING DATE: 2001-02-02  
;; PRIOR APPLICATION NUMBER: PCT/US01/03537  
;; PRIORITY FILING DATE: 2001-02-02  
;; PRIOR APPLICATION NUMBER: 09/497,625  
;; PRIORITY FILING DATE: 2000-02-03  
;; PRIOR APPLICATION NUMBER: 09/359,193  
;; PRIORITY FILING DATE: 1999-07-22  
;; PRIOR APPLICATION NUMBER: 09/121,781  
;; PRIORITY FILING DATE: 1998-07-23  
;; NUMBER OF SEQ ID NOS: 107  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 23  
;; LENGTH: 100  
;; TYPE: PRT  
;; ORGANISM: Mus musculus  
US-09-840-459-23

Query Match 79.0%; Score 64; DB 10; Length 100;  
Best Local Similarity 75.0%; Pred. No. 0.0017;  
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||:|||||  
Db 24 KSSQSLLYSNGKTYLN 39

RESULT 7

US-09-840-459-24  
; Sequence 24, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: LaRosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Siobhan H.  
; APPLICANT: O'Keefe, Theresa

;; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
;; FILE REFERENCE: 1855.1052-012  
;; CURRENT APPLICATION NUMBER: US/09/840,459  
;; PRIORITY FILING DATE: 2001-02-02  
;; PRIOR APPLICATION NUMBER: PCT/US01/03537  
;; PRIORITY FILING DATE: 2001-02-02  
;; PRIOR APPLICATION NUMBER: 09/497,625  
;; PRIORITY FILING DATE: 2000-02-03  
;; PRIOR APPLICATION NUMBER: 09/359,193  
;; PRIORITY FILING DATE: 1999-07-22  
;; PRIOR APPLICATION NUMBER: 09/121,781  
;; PRIORITY FILING DATE: 1998-07-23  
;; NUMBER OF SEQ ID NOS: 107  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 24  
;; LENGTH: 100  
;; TYPE: PRT  
;; ORGANISM: Mus musculus  
;; NAME/KEY: VARIANT  
;; LOCATION: (1)...(100)  
;; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-840-459-24

Query Match 79.0%; Score 64; DB 10; Length 100;  
Best Local Similarity 75.0%; Pred. No. 0.0017;  
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||:|||||  
Db 24 KSSQSLLYSNGKTYLN 39

RESULT 8

US-09-840-459-56  
; Sequence 56, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: LaRosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Siobhan H.  
; APPLICANT: O'Keefe, Theresa

;; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
;; FILE REFERENCE: 1855.1052-012  
;; CURRENT APPLICATION NUMBER: US/09/840,459  
;; PRIORITY FILING DATE: 2001-02-02  
;; PRIOR APPLICATION NUMBER: PCT/US01/03537  
;; PRIORITY FILING DATE: 2001-02-02  
;; PRIOR APPLICATION NUMBER: 09/497,625  
;; PRIORITY FILING DATE: 2000-02-03  
;; PRIOR APPLICATION NUMBER: 09/359,193  
;; PRIORITY FILING DATE: 1999-07-22  
;; PRIOR APPLICATION NUMBER: 09/121,781  
;; PRIORITY FILING DATE: 1998-07-23  
;; NUMBER OF SEQ ID NOS: 107  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 56  
;; LENGTH: 112  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-840-459-56

Query Match 79.0%; Score 64; DB 10; Length 112;  
Best Local Similarity 75.0%; Pred. No. 0.0019;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||:|||||  
Db 24 RSSQSLVYSDGNTYLN 39

RESULT 9

US-09-840-459-66  
; Sequence 66, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: LaRosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Siobhan H.  
; APPLICANT: O'Keefe, Theresa

;; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
;; FILE REFERENCE: 1855.1052-012  
;; CURRENT APPLICATION NUMBER: US/09/840,459  
;; PRIORITY FILING DATE: 2001-02-02  
;; PRIOR APPLICATION NUMBER: PCT/US01/03537  
;; PRIORITY FILING DATE: 2001-02-02  
;; PRIOR APPLICATION NUMBER: 09/497,625  
;; PRIORITY FILING DATE: 2000-02-03  
;; PRIOR APPLICATION NUMBER: 09/359,193  
;; PRIORITY FILING DATE: 1999-07-22  
;; PRIOR APPLICATION NUMBER: 09/121,781  
;; PRIORITY FILING DATE: 1998-07-23  
;; NUMBER OF SEQ ID NOS: 107  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 66  
;; LENGTH: 112  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

Query Match 79.0%; Score 64; DB 10; Length 100;  
Best Local Similarity 75.0%; Pred. No. 0.0017;  
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||:|||||  
Db 24 KSSQSLLYSNGKTYLN 39

US-09-840-459-66

Query Match 79.0%; Score 64; DB 10; Length 112;  
Best Local Similarity 75.0%; Pred. No. 0.0019;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSSKSLLYKDGKTYLN 16  
|||:|:|:|:|:|:|  
Db 24 RSSQSLVYSDGNTYLN 39

## RESULT 10

```

US-10-006-773-11
; Sequence 11, Application US/10006773
; Patent No. US20020132983A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Ant
; FILE REFERENCE: 003
; CURRENT APPLICATION NUMBER: US/10/006,773
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/250,089
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-006-773-11

```

**Qy**            1 R S S K S L L Y K D G K T Y L N     16  
              : | | : | | | | : | | | | |  
**Dd**            43 K S S Q S L L Y S N G K T Y L N     58

RESULT 11

```

US-09-479-614-26
; Sequence 26, Application US/09479614
; Publication NO. US2003001383A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Felis catus
; US-09-479-614-26

```

Qy            1 R S S K S L L Y K D G K T Y L N     16  
             | : | : | : | : | : | :  
Db          24 R A S O S L L Y S D G N T Y L N     39

13. T. Williams

US-09-479-614-20  
US-09-479-614-20  
; Sequence 20; Application US/09479614

Qy 1 RSSKSLLYKDGKTYLN 16  
|||:::|||||  
Db 24 RSSOSLVFSDGNTYLN 39



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:46:53 ; Search time 3.5 Seconds  
(without alignments)  
82.953 Million cell updates/sec

Title: US-09-674-716b-5

Perfect score: 31

Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	31	100.0	112	1 KV2A MOUSE	P01626 mus musculus
2	31	100.0	113	1 KV2C MOUSE	P01628 mus musculus
3	31	100.0	120	1 KV2E MOUSE	P01627 mus musculus
4	27	87.1	104	1 YHGB BACSU	P38048 bacillus su
5	27	87.1	234	1 YHKO YEAST	P40011 saccharomyc
6	27	87.1	760	1 MCM7 SCHPO	O75001 schizosacch
7	27	87.1	845	1 CC47 YEAST	P38132 saccharomyc
8	27	87.1	1401	1 WIS4 SCHPO	O14299 schizosacch
9	26	83.9	193	1 RS3A SULTO	Q97518 sulfolobus
10	26	83.9	341	1 NU2M DROME	P03896 drosophila
11	26	83.9	341	1 NU2M DROYA	P03895 drosophila
12	26	83.9	1453	1 CALL1 MOUSE	P11087 mus musculus
13	26	83.9	1460	1 CALL1 CANFA	O92617 canis famil
14	26	83.9	1464	1 CALL1 HUMAN	P02452 homo sapien
15	25	80.6	212	1 AG11 WHEAT	P10968 triticum ae
16	25	80.6	212	1 AG1 HORVU	P15312 hordeum vul
17	25	80.6	463	1 D2DR FUGRU	P53453 fugu rubrip
18	25	80.6	565	1 TP6B AERPE	Q9ye64 aeropyrum p
19	25	80.6	664	1 ALOX PTCAN	P04841 pichia angu
20	25	80.6	778	1 YFK8 YEAST	P43610 saccharomyc
21	25	80.6	887	1 YAY5 SCHPO	Q10213 schizosacch
22	24	77.4	176	1 Y059 NPVOP	O10316 orgyia pseu
23	24	77.4	228	1 EBP CAVPO	Q60490 cavia porce
24	24	77.4	230	1 EBP MOUSE	P70245 mus musculus
25	24	77.4	250	1 PP2C BRUME	Q57099 bruceella me
26	24	77.4	259	1 PP26 BRUAB	Q4642 bruceella ab
27	24	77.4	312	1 IG37 YEAST	P53275 saccharomyc
28	24	77.4	339	1 G3P MYCTU	O06822 mycobacteri
29	24	77.4	443	1 FBL4 CRIGR	O55058 cricetus
30	24	77.4	443	1 FBL4 HUMAN	O59667 homo sapien
31	24	77.4	443	1 FBL4 MOUSE	Q9wv79 mus muscul
32	24	77.4	463	1 SAHH CAUCR	O9ab10 caulobacter
33	24	77.4	543	1 NFL_HUMAN	P07196 homo sapien

RESULT 1

ID	KV2A MOUSE	STANDARD;	PRT;	112 AA.
AC	P01626;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG kappa chain V-II region MOPC 167.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=79000273; PubMed=99160;			
RA	Rudikoff S., Potter M.;			
RT	"Kappa Chain variable region from M167, a phosphorylcholine binding			
RL	(Biochemistry 17:2703-2707(1978)).			
CC	-1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT			
CC	BINDS PHOSPHORYLCHOLINE. THE SEQUENCE OF THE V REGION OF THE HEAVY			
CC	CHAIN HAS ALSO BEEN DETERMINED.			
DR	PIR; A01908; KVM516.			
DR	HSP; P80362; IWL.			
DR	InterPro; IPR003006; IG_MHC.			
DR	InterPro; IPR003596; IG_V.			
DR	Fram; PF00047; IG; 1.			
DR	SMART; SM00406; IGV; 1.			
KW	immunoglobulin V region.			
FT	DOMAIN 1 23			
FT	DOMAIN 24 39			
FT	DOMAIN 40 54			
FT	DOMAIN 55 61			
FT	DOMAIN 62 93			
FT	DOMAIN 94 102			
FT	DOMAIN 103 112			
FT	DISULFID 123 93			
FT	NON_TER 112 112			
SQ	SEQUENCE 112 AA; 12349 MW; A58EDFD6404B9726 CRC64;			

Query Match 100.0%; Score 31; DB 1; Length 112;

Best Local Similarity 100.0%; Pred. No. 0.29;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

DB 55 LMSTRAS 61

RESULT 2

ID	KV2C MOUSE	STANDARD;	PRT;	113 AA.
AC	P01628;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			

```

DE IG kappa chain V-II region MOPC 511.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=81052016; PubMed=6776396;
RA Appella E.;
RT "Amino acid sequence of the light chain variable region of M511, a
RT phosphorylcholine-binding murine myeloma protein.";
RL Mol. Immunol. 17:711-718(1980).
CC -1- MISCELLANEOUS; THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS PHOSPHORYLCHOLINE.
CC PIR; A01910; KVM51.
DR HSP; P80362; IWTI.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NCBI_TaxID=10090;
KW Immunoglobulin V region.
FT SIGNAL 1 23
FT CHAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DISULFID 23 93
FT NON TER 113 113
SQ SEQUENCE 113 AA; 12496 MW; EFB0DC4DA2BD3450 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
DB 55 LMSTRAS 61

RESULT 3
KV2B_MOUSE STANDARD; PRT; 120 AA.
AC P01627;
AT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG kappa chain V-II region VKAPPA167 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82002223; PubMed=6791832;
RA Seising E., Storb U.;
RT "Somatic mutation of immunoglobulin light-chain variable-region
RT genes.";
RL Cell 25:47-58(1981).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; J00562; AAA39032.1; -
DR AMBL; K02415; AAA39051.1; -
DR PIR; A01909; KVM567.
DR HSP; P80362; IWTI.

```

```

DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 120
FT DOMAIN 21 43
FT DOMAIN 44 59
FT DOMAIN 60 74
FT DOMAIN 75 81
FT DOMAIN 82 113
FT DOMAIN 114 120
FT DISULFID 43 113
FT SEQUENCE 120 AA; 13280 MW; 63BB571F0E4DE3E8 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
DB 75 LMSTRAS 81

RESULT 4
YHGB_BACSU STANDARD; PRT; 104 AA.
ID YHGB_BACSU
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yhgb.
GN YHGB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Noback M.A., Terpestra P., Holsappel S., Venema G., Bron S.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunat F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritze C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parrean E., Pucic P., Purnelle D., Porwollik S., Prescott A.M.,
RA Riegler M., Rivolta C., Rocha E., Roche S., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield P.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitznegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

```

RA Yoshida K., Yoshikawa H.F., Zunstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
RT *subtilis*.";  
RL Nature 390:249-256(1997).  
RN [3]  
RP SEQUENCE OF 1-67 FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=93328693; PubMed=8335642;  
RA Popham D.L., Setlow P.;  
RT "Cloning, nucleotide sequence, and regulation of the *Bacillus*  
RT *subtilis* pbpF gene, which codes for a putative class A  
RT high-molecular-weight penicillin-binding protein.";  
RL J. Bacteriol. 175:4870-4876(1993).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Y14083; CAA74515.1; -;  
DR EMBL; Z99109; CAB12849.1; -;  
DR EMBL; L10630; AAA71940.2; -;  
DR PIR; C40614; C40614.  
DR Subtilist; BG10426; yhgB.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 104 AA; 11998 MW; 6B4CD77B52FD968 CRC64;  
  
Query Match 87.1%; Score 27; DB 1; Length 104;  
Best Local Similarity 85.7%; Pred. No. 3.1;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 LMSTRAS 7  
Db 38 LMSTRCS 44  
  
RESULT 5  
YEKO\_YEAST  
ID YEKO\_YEAST STANDARD; PRT; 234 AA.  
AC P40011;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Hypothetical 25.6 kDa protein in NTF2-SRPI intergenic region.  
GN YER010C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / ABU72;  
RA Dietrich F.S., Mullen J.T., Hennessey K.M., Allen E., Araujo R.,  
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,  
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U18778; AAB64543.1; -;  
DR

DR SGD; S0000812; YER010C.  
KW Hypothetical protein.  
SQ SEQUENCE 234 AA; 25563 MW; EC109F224240F980 CRC64;  
  
Query Match 87.1%; Score 27; DB 1; Length 234;  
Best Local Similarity 100.0%; Pred. No. 7.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 LMSTRA 6  
Db 102 LMSTRA 107  
  
RESULT 6  
MCM7\_SCHPO  
ID MCM7\_SCHPO STANDARD; PRT; 760 AA.  
AC O75001; P87302;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE DNA replication licensing factor mcm7 (Minichromosome maintenance  
DE protein 7).  
GN MCM7 OR SPBC25D12.03C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=011;  
RA Liang D.T., Forsburg S.L.;  
RT "Fission yeast mcm7+ is an essential gene required for normal DNA  
RT replication";  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagals K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Mosel D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas R., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of *Schizosaccharomyces pombe*."  
RL Nature 415:871-880(2002).  
RN [3]  
RP SEQUENCE OF 367-466 FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=98031880; PubMed=9366552;  
RA Adachi Y., Usukura J., Yanagida M.;  
RT "A globular complex formation by Nda1 and the other five members of  
RT the MCM protein family in fission yeast."  
RL Genes Cells 2:467-479(1997).

```

CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF070481; AAC33693.1; -
CC EMBL; AL031158; CAA20099.1; -
CC EMBL; AJ000065; CAA03898.1; -
CC InterPro; IPR001208; MCM.
CC Pfam; PF00493; MCM; 1.
CC ProDom; PD001041; MCM; 1.
CC SMART; SM00350; MCM; 1.
CC PROSITE; PS00847; MCM 1; 1.
CC PROSITE; PS0051; MCM 2; 1.
CC Transcription regulation; DNA-binding; Nuclear protein;
CC DNA replication; Cell cycle; ATP-binding.
CC DOMAIN 353 559 MCM.
CC NP_BIND 403 410 ATP (POTENTIAL).
CC NP_BIND 457 457 V -> I (IN REF. 3).
CC CONFLICT 457 457
CC SEQUENCE 760 AA; 85622 MW; 2141F8F9CB0BAA34 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MSTRAS 7
DB 243 MSTRAS 248

RESULT 7
CC47_YEAST STANDARD; PRT; 845 AA.
ID CC47_YEAST
AC F8132;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA replication licensing factor CDC47 (Cell division control protein
DE 47).
DE CDC47 OR YER202W OR YER1441.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Dalton S.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=93377417; PubMed=8368014;
RA Bussereau F., Mallet L., Gallion L., Jacquet M.;
RT "A 12.8 kb segment, on the right arm of chromosome II from
RT Saccharomyces cerevisiae including part of the DURL2 gene, contains
RT five putative new genes.";
RL Yeast 9:797-806 (1993).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; U14730; AAA86309.1; -
CC EMBL; Z21487; CAA79689.1; -
CC EMBL; Z36071; CAA85166.1; -
CC PIR; S34027; S34027.
CC SGD; S0000406; CDC47.
CC InterPro; IPR001208; MCM.
CC Pfam; PF00493; MCM; 1.
CC ProDom; PD001041; MCM; 1.
CC SMART; SM00350; MCM; 1.
CC PROSITE; PS00847; MCM 1; 1.
CC PROSITE; PS0051; MCM 2; 1.
CC Transcription regulation; DNA-binding; Nuclear protein;
CC DNA replication; Cell cycle; ATP-binding.
CC DOMAIN 410 617 MCM.
CC NP_BIND 460 467 ATP (POTENTIAL).
CC CONFLICT 552 552 G -> V (IN REF. 2).
CC CONFLICT 556 558 TLN -> NFG (IN REF. 2).
CC CONFLICT 574 574 Y -> I (IN REF. 2).
CC SEQUENCE 845 AA; 94942 MW; ADA66C719D96DB4A CRC64;

Query Match 87.1%; Score 27; DB 1; Length 845;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MSTRAS 7
DB 300 MSTRAS 305

RESULT 8
WIS4_SCHPO STANDARD; PRT; 1401 AA.
ID WIS4_SCHPO
AC O14299; P87062; Q92384;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE MAP kinase kinase kinase wis4 (EC 2.7.1.-) (MAP kinase kinase kinase
DE wisk1) (MAP kinase kinase kinase wisk1).
DE WIS4 OR WAK1 OR WIK1 OR SPAC9G1.02.
GN Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98026892; PubMed=9321395;
RA Samejima I., Mackie S., Fantes P.A.;
RT "Multiple modes of activation of the stress-responsive MAP kinase
RT pathway in fission yeast.";
RL EMBO J. 16:6162-6170 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh J.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkstraet G., Aert K., Robben J., Grymprez B.,
RA Welljens I., Vanstraelen E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Fritz C., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

```



RA Goffeau A., Cadieu B., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of *Schizosaccharomyces pombe*.";  
 RL Nature 415:871-880(2002).  
 RN [3]  
 RP SEQUENCE OF 96-1401 FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=97282620; PubMed=9136929;  
 RA Shieh J.C., Wilkinson M., Buck V., Morgan B., Makino K.,  
 RA Millar J.B.A.;  
 RT "The Mca4 response regulator coordinately controls the stress-  
 RT activated Wsk1-Wsk1-Styl MAP kinase pathway and fission yeast cell  
 RT cycle.";  
 RL Genes Dev. 11:1008-1022(1997).  
 RN [4]  
 RP SEQUENCE OF 457-543 FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=973131762; PubMed=9188094;  
 RA Shiozaki K., Shiozaki M., Russell P.;  
 RT "Mca4 mitotic catastrophe suppressor regulates the fission yeast cell  
 RT cycle through the Wsk1-Wsk1-Spcl kinase cascade.";  
 RL Mol. Biol. Cell 8:409-419(1997).  
 CC -!- FUNCTION: INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT IS  
 CC ACTIVATED IN UNDER CONDITIONS OF HEAT SHOCK, OXIDATIVE STRESS OR  
 CC LIMITED NUTRITION. UNLIKE WSK1, IT IS NOT ACTIVATED BY CHANGES IN  
 CC THE OSMOLARITY OF THE EXTRACELLULAR ENVIRONMENT. ACTIVATES THE  
 CC WSK1 MAP KINASE KINASE BY PHOSPHORYLATION.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z98763; CAB11500.1; -;  
 CC EMBL; Y07750; CAA69030.1; ALT\_INIT.  
 CC EMBL; Y11989; CAA72718.1; -;  
 CC EMBL; U81521; AAB39762.1; -;  
 CC HSP; P24941; 1CXP.  
 CC InterPro; IPR000719; Euk\_pkinase.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC Pfam; PF00069; pkinase; 1.  
 CC ProDom; PD000001; Euk\_pkinase; 1.  
 CC SMART; SM00220; S\_TKc; 1.  
 CC PROSITE; PS00107; PROTEIN KINASE ATP; FALSE\_NEG.  
 CC PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
 CC PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
 FT DOMAIN 1037 1306 PROTEIN KINASE.  
 FT NP\_BIND 1043 1051 ATP (BY SIMILARITY).  
 FT BINDING 1066 1066 ATP (BY SIMILARITY).  
 FT ACT\_SITE 1161 1161 BY SIMILARITY.  
 FT CONFLICT 484 485 RL -> SP (IN REF. 3).  
 FT CONFLICT 1346 1346 D -> V (IN REF. 1).  
 SQ SEQUENCE 1401 AA; 160537 MW; 529A989D2B627F3D CRC64;

Query Match 87.1%; Score 27; DB 1; Length 1401;  
 Best Local Similarity 85.7%; Pred. No. 58;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LMSTRAS 7  
 DB 81 LMSTRAS 87

RESULT 9  
 RS3A\_SULTO STANDARD; PRT; 193 AA.  
 ID RS3A\_SULTO  
 AC Q975F8;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 30S ribosomal protein S3Ae.  
 GN RPS3AE OR S70453.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OX NCBI\_TaxID=111955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JCM 10545 / 7;  
 RX MEDLINE=21456156; PubMed=11572479;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,  
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
 RA Oshima T., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic thermophilic  
 RT Crenarchaeon, *Sulfolobus tokodaii* strain 7.";  
 RL DNA Res. 8:123-140(2001).  
 CC -!- SIMILARITY: BELONGS TO THE S3AE FAMILY OF RIBOSOMAL PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AP000982; BAB65443.1; -;  
 CC InterPro; IPR001593; Ribosomal\_S3AE.  
 CC Pfam; PF01015; Ribosomal\_S3AE; 1.  
 CC ProDom; PD003035; Ribosomal\_S3AE; 1.  
 CC PROSITE; PS01191; RIBOSOMAL\_S3AE; FALSE\_NEG.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 193 AA; 22180 MW; 78A94B87B961D751 CRC64;  
 Query Match 83.9%; Score 26; DB 1; Length 193;  
 Best Local Similarity 71.4%; Pred. No. 12;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LMSTRAS 7  
 DB 142 ILSTRAS 148

RESULT 10  
 NU2M\_DROME STANDARD; PRT; 341 AA.  
 ID NU2M\_DROME  
 AC P03896;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 2 (EC 1.6.5.3).  
 GN ND2.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Mitochondrion.  
 OC Rukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Sphingomyia; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96423163; PubMed=8825764;  
 RA Lewis D.L., Farr C.L., Kaguni L.S.;

```

RT "Drosophila melanogaster mitochondrial DNA: completion of the
RT nucleotide sequence and evolutionary Comparisons.";
RL Insect Mol. Biol. 4:263-278(1995).
RN [2]
RX SEQUENCE OF 56-341 FROM N.A.
RP MEDLINE=83245048; PubMed=6408489;
RA de Bruijn M.H.L.;
RT "Drosophila melanogaster mitochondrial DNA, a novel organization and
RT genetic code.";
RL Nature 304:234-241(1983).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U37541; AAC47811.1; -.
DR EMBL; J01404; AAB59238.1; -.
DR PIR; A00419; OXFF2M.
DR FlyBase; FBgn0013680; mt:ND2.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
KW OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion; Transmembrane.
SQ SEQUENCE 341 AA; 39773 MW; 2B8E9528C28007D8 CRC64;

Query Match 83.9%; Score 26; DB 1; Length 341;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 51 LMSTEAS 57

RESULT 11
NU2M DROVA
ID NU2M DROVA STANDARD; PRT; 341 AA.
AC P03895;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 2 [EC 1.6.5.3].
GN ND2.
OS Drosophila yakuba (Fruit fly).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7245;
[1]
RX SEQUENCE FROM N.A.
RP STRAIN=2317 6 Ivory Coast;
RC MEDLINE=86089137; PubMed=3001325;
RA Clary D.O., Wolstenholme D.R.;
RT "The mitochondrial DNA molecular of Drosophila yakuba: nucleotide
RT sequence, gene organization, and genetic code.";
RL J. Mol. Evol. 22:252-271(1985).
RN [2]
RX SEQUENCE FROM N.A.
RP MEDLINE=84041489; PubMed=6314262;
RA Clary D.O., Wolstenholme D.R.;
RT "Genes for cytochrome c oxidase subunit I, URF2, and three tRNAs in
RT Drosophila mitochondrial DNA.";
RL Nucleic Acids Res. 11:6859-6872(1983).
RN [3]
RX SEQUENCE OF 1-56 FROM N.A.

```

```

RX MEDLINE=83090428; PubMed=6294611;
RA Clary D.O., Goddard J.M., Martin S.C., Fauron C.M.-R.,
RA Wolstenholme D.R.;
RT "Drosophila mitochondrial DNA: a novel gene order.";
RL Nucleic Acids Res. 10:6619-6637(1982).
RN [4]
RX SEQUENCE OF 1-56 FROM N.A.
RP MEDLINE=88011348; PubMed=3116271;
RA Clary D.O., Wolstenholme D.R.;
RT "Drosophila mitochondrial DNA: conserved sequences in the A + T-rich
RT region and supporting evidence for a secondary structure model of the
RT small ribosomal RNA.";
RL J. Mol. Evol. 25:116-125(1987).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03240; CAA26985.1; -.
DR EMBL; X05915; CAA29342.1; -.
DR PIR; A00418; OXFF2Y.
DR PIR; A25797; A25797.
DR FlyBase; FBgn0013184; Dyak\mt:ND2.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
KW OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion; Transmembrane.
SQ SEQUENCE 341 AA; 39495 MW; E6117DE508E9D4AF CRC64;

Query Match 83.9%; Score 26; DB 1; Length 341;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 51 LMSTEAS 57

RESULT 12
CALL MOUSE
ID CALL MOUSE STANDARD; PRT; 1453 AA.
AC P11087; Q60635;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(I) chain precursor.
DE COL1A1 OR COLA1.
GN COL1A1 OR COLA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RX SEQUENCE FROM N.A.
RP STRAIN=FVB/N;
RC MEDLINE=96033240; PubMed=8535610;
RA Li S.W., Khillan J., Prockop D.J.;
RT "The complete cDNA coding sequence for the mouse pro alpha 1(I) chain
RT of type I procollagen.";
RL Matrix Biol. 14:593-595(1995).
RN [2]
RX SEQUENCE OF 518-1128 FROM N.A.
RP MEDLINE=86137403; PubMed=3841523;
RA French B.T., Lee W.-H., Maul G.G.;
RT "Nucleotide sequence of a cDNA clone for mouse pro alpha 1(I)
RT collagen protein.";
RL Gene 39:311-312(1985).

```

[3] RP SEQUENCE OF 735-878 AND 1005-1058 FROM N.A.  
 RX MEDLINE=831511374; PubMed=6298597;  
 RA Monson J.M., Friedman J., McCarthy B.J.;  
 RT "DNA sequence analysis of a mouse pro alpha 1 (I) procollagen gene:  
 FT evidence for a mouse B1 element within the gene";  
 RL Mol. Cell. Biol. 2:1362-1371(1982).  
 [4] RP SEQUENCE OF 735-878 AND 1005-1058 FROM N.A.  
 RX MEDLINE=83151109; PubMed=6219867;  
 RA Monson J.M., McCarthy B.J.;  
 RT "Identification of a Balb/c mouse pro alpha 1(I) procollagen gene:  
 FT evidence for insertions or deletions in gene coding sequences";  
 RL DNA 1:59-69(1981).  
 [5] RP SEQUENCE OF 1442-1453 FROM N.A.  
 RX MEDLINE=88124276; PubMed=3340560;  
 RA Mooslechner K., Harbers K.;  
 RT "Two mRNAs of mouse pro alpha 1(I) collagen gene differ in the size  
 of the 3'-untranslated region";  
 RL Nucleic Acids Res. 16:773-773(1988).  
 CC -!- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN  
 (FIBRILLAR FORMING COLLAGEN).  
 CC -!- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.  
 CC -!- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND  
 BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM  
 HYDROXYAPATITE.  
 CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING  
 UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.  
 CC -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U08020; AAA88912.1; -;  
 DR EMBL; X15896; CAA33904.1; -;  
 DR EMBL; M14423; AAA37333.1; -;  
 DR EMBL; M17491; AAA37334.1; -;  
 DR EMBL; X06753; CAA29927.1; -;  
 DR EMBL; K03036; AAA37332.1; -;  
 DR EMBL; K03029; AAA37332.1; JOINED.  
 DR EMBL; K03030; AAA37332.1; JOINED.  
 DR EMBL; K03031; AAA37332.1; JOINED.  
 DR EMBL; K03032; AAA37332.1; JOINED.  
 DR EMBL; K03033; AAA37332.1; JOINED.  
 DR EMBL; K03034; AAA37332.1; JOINED.  
 DR EMBL; K03035; AAA37332.1; JOINED.  
 DR PIR; A23982; A23982.  
 DR MGD; MGI:88467; Col1a1.  
 DR InterPro; IPR000087; Collagen.  
 DR InterPro; IPR000885; Fib collagen\_C.  
 DR InterPro; IPR001007; VWFC.  
 DR Pfam; PF01391; Collagen\_18.  
 DR Pfam; PF01410; COLFI; 1.  
 DR ProDom; PD000007; Collagen; 1.  
 DR ProDom; PD002078; Fib collagen\_C; 1.  
 DR SMART; SM00038; COLFI; 1.  
 DR SMART; SM00214; VWFC; 1.  
 DR PROSITE; PS01208; VWFC; 1.  
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;  
 KW Glycoprotein; Collagen; Signal.  
 FT SIGNAL 1 22  
 FT PROPEP 23 151 AMINO-TERMINAL PROPEPTIDE.  
 FT CHAIN 152 1207 COLLAGEN ALPHA 1(I) CHAIN.  
 FT PROPEP 1208 1453 CARBOXYL-TERMINAL PROPEPTIDE.  
 FT DOMAIN 29 87 VWFC.  
 FT DOMAIN 152 167 NONHELICAL REGION (N-TERMINAL).  
 FT DOMAIN 168 1181 TRIPLE-HELICAL REGION.

FT DOMAIN 1182 1207 NONHELICAL REGION (C-TERMINAL).  
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1354 1354 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 734 736 CELL ATTACHMENT SITE (POTENTIAL).  
 FT SITE 1082 1084 CELL ATTACHMENT SITE (POTENTIAL).  
 FT CONFLICT 1450 1450 A -> V (IN REF. 5).  
 SQ SEQUENCE 1453 AA; 137944 MW; 3B802E535DF81808 CRC64;  
 Query Match 83.9%; Score 26; DB 1; Length 1453;  
 Best Local Similarity 85.7%; Pred. No. 1.le-02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LMSTRAS 7  
 Db 1346 LMSTRAS 1352  
 RESULT 13  
 ID CANFA  
 ID CANFA STANDARD; PRT; 1460 AA.  
 AC Q9XSJ7.  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Collagen alpha 1(I) chain precursor.  
 GN COL1A1.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Skin;  
 RA Campbell B.G., Wootton J.A.M., McLeod J.N., Minor R.R.;  
 RT "Sequence of normal canine COL1A1 cDNA";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN  
 (FIBRILLAR FORMING COLLAGEN).  
 CC -!- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.  
 CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING  
 UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.  
 CC -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF153062; AAD34619.1; -;  
 DR InterPro; IPR000087; Collagen.  
 DR InterPro; IPR000885; Fib collagen\_C.  
 DR InterPro; IPR001007; VWFC.  
 DR Pfam; PF01391; Collagen\_18.  
 DR Pfam; PF01410; COLFI; 1.  
 DR ProDom; PD000007; Collagen; 1.  
 DR ProDom; PD002078; Fib collagen\_C; 1.  
 DR SMART; SM00038; COLFI; 1.  
 DR SMART; SM00214; VWFC; 1.  
 DR PROSITE; PS01208; VWFC; 1.  
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;  
 KW Glycoprotein; Collagen; Signal.  
 FT SIGNAL 1 22  
 FT PROPEP 23 157 AMINO-TERMINAL PROPEPTIDE.  
 FT CHAIN 158 1214 COLLAGEN ALPHA 1(I) CHAIN.  
 FT PROPEP 1215 1460 CARBOXYL-TERMINAL PROPEPTIDE.  
 FT DOMAIN 34 52 VWFC.  
 FT DOMAIN 158 174 NONHELICAL REGION (N-TERMINAL).  
 FT DOMAIN 175 1188 TRIPLE-HELICAL REGION.  
 FT DOMAIN 1189 1214 NONHELICAL REGION (C-TERMINAL).  
 FT DOMAIN 741 743 CELL ATTACHMENT SITE (POTENTIAL).

FT SITE 1089 1091 CELL ATTACHMENT SITE (POTENTIAL).  
 FT CARBOHYD 1361 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 1460 AA; 138762 NW; 58E3674D2B570697 CRC64;  
 Query Match 83.9%; Score 26; DB 1; Length 1460;  
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LMSTRAS 7  
 |||||  
 Db 1353 LMSTRAS 1359

RESULT 14  
 CALL\_HUMAN  
 ID CALL\_HUMAN STANDARD; PRT; 1464 AA.  
 AC P02452; Q15176; Q14037;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Collagen alpha 1(I) chain precursor.  
 GN COL1A1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 1-472 FROM N.A.  
 RX MEDLINE=89025644; PubMed=3178743;  
 RA Tromp G., Kuivaniemi H., Stacey A., Shikata H., Baldwin C.T.,  
 Jaenisch R., Prockop D.J.;  
 RT "Structure of a full-length cDNA clone for the prepro alpha 1(I)  
 chain of human type I procollagen.";  
 RL Biochem. J. 253:919-922(1988).  
 RN [2]  
 RP SEQUENCE OF 1-181 FROM N.A.  
 RX MEDLINE=84270697; PubMed=6462220;  
 RA Chu M.-L., de Wet W.J., Bernard M.P., Ding J.-F., Morabito M.,  
 Myers J., Williams C., Ramirez F.;  
 RT "Human pro alpha 1(I) collagen gene structure reveals evolutionary  
 conservation of a pattern of introns and exons.";  
 RL Nature 310:337-340(1984).  
 RN [3]  
 RP SEQUENCE OF 162-301.  
 RC TISSUE=Skin;  
 RX MEDLINE=71038625; PubMed=5529814;  
 RA Click E.M., Bornstein P.;  
 RT "Isolation and characterization of the cyanogen bromide peptides from  
 the alpha 1 and alpha 2 chains of human skin collagen.";  
 RL Biochemistry 9:4699-4706(1970).  
 RN [4]  
 RP SEQUENCE OF 263-268.  
 RC TISSUE=Skin;  
 RX MEDLINE=71001508; PubMed=4319110;  
 RA Morgan P.H., Jacobs H.G., Segrest J.P., Cunningham L.W.;  
 RT "A comparative study of glycopeptides derived from selected  
 vertebrate collagens. A possible role of the carbohydrate in fibril  
 formation.";  
 RL J. Biol. Chem. 245:5042-5048(1970).  
 RN [5]  
 RP SEQUENCE OF 425-1464 FROM N.A.  
 RX MEDLINE=84080385; PubMed=6689127;  
 RA Bernard M.P., Chu M.-L., Myers J.C., Ramirez F., Eikenberry E.F.,  
 Prockop D.J.;  
 RT "Nucleotide sequences of complementary deoxyribonucleic acids for the  
 pro alpha 1 chain of human type I procollagen. Statistical evaluation  
 of structures that are conserved during evolution.";  
 RL Biochemistry 22:5213-5223(1983).  
 RN [6]  
 RP SEQUENCE OF 1229-1454 FROM N.A.  
 RC TISSUE=Bone;  
 RX MEDLINE=88124208; PubMed=3340531;  
 RA Maekelae J.K., Raassina M., Virta A., Vuorio E.;

RT "Human pro alpha 1(I) collagen: cDNA sequence for the C-propeptide  
 domain.";  
 RL Nucleic Acids Res. 16:349-349(1988).  
 RN [7]  
 RP SEQUENCE OF 1-34 FROM N.A.  
 RX MEDLINE=88097389; PubMed=3480516;  
 RA Bornstein P., McKay J., Morishima J.K., Devarayalu S., Gelinas R.E.;  
 RT "Regulatory elements in the first intron contribute to  
 transcriptional control of the human alpha 1(I) collagen gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8869-8873(1987).  
 RN [8]  
 RP SEQUENCE OF 1-34 FROM N.A.  
 RX MEDLINE=85130970; PubMed=2857713;  
 RA Chu M.-L., de Wet W.J., Bernard M.P., Ramirez F.;  
 RT "Fine structural analysis of the human pro-alpha 1 (I) collagen gene.  
 Promoter structure, Alu repeats, and polymorphic transcripts.";  
 RL J. Biol. Chem. 260:2315-2320(1985).  
 RN [9]  
 RP SEQUENCE OF 1-44 FROM N.A.  
 RX MEDLINE=88033098; PubMed=2822714;  
 RA Rossouw C.M.S., Vergeer W.P., du Plooy S.J., Bernard M.P., Ramirez F.,  
 de Wet W.J.;  
 RT "DNA sequences in the first intron of the human pro-alpha 1(I)  
 collagen gene enhance transcription.";  
 RL J. Biol. Chem. 262:15151-15157(1987).  
 RN [10]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=91184577; PubMed=2010058;  
 RA Kuivaniemi H., Tromp G., Prockop D.J.;  
 RT "Mutations in collagen genes: causes of rare and some common diseases  
 in humans.";  
 RL FASEB J. 5:2052-2060(1991).  
 RN [11]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=97255959; PubMed=9101290;  
 RA Kuivaniemi H., Tromp G., Prockop D.J.;  
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-  
 associated collagen (type IX), and network-forming collagen (type X)  
 cause a spectrum of diseases of bone, cartilage, and blood vessels.";  
 RL Hum. Mutat. 9:300-315(1997).  
 RN [12]  
 RP REVIEW ON OI VARIANTS.  
 RX MEDLINE=91374476; PubMed=1895312;  
 RA Byers P.H., Wallis G.A., Willing M.C.;  
 RT "Osteogenesis imperfecta: translation of mutation to phenotype.";  
 RL J. Med. Genet. 28:433-442(1991).  
 RN [13]  
 RP REVIEW ON OI VARIANTS.  
 RX MEDLINE=97169389; PubMed=9016532;  
 RA Dalgleish R.;  
 RT "The human type I collagen mutation database.";  
 RL Nucleic Acids Res. 25:181-187(1997).  
 RN [14]  
 RP VARIANT OI-II CYS-1166.  
 RX MEDLINE=86287390; PubMed=3016737;  
 RA Cohn D.H., Byers P.H., Steinmann B., Gelinas R.E.;  
 RT "Lethal osteogenesis imperfecta resulting from a single nucleotide  
 change in one human pro alpha 1(I) collagen allele.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6045-6047(1986).  
 RN [15]  
 RP VARIANT OI-II ARG-569.  
 RX MEDLINE=87222295; PubMed=3108247;  
 RA Bateman J.F., Chan D., Walkers I.D., Rogers J.G., Cole W.G.;  
 RT "Lethal perinatal osteogenesis imperfecta due to the substitution of  
 arginine for glycine at residue 391 of the alpha 1(I) chain of type I  
 collagen.";  
 RL J. Biol. Chem. 262:7021-7027(1987).  
 RN [16]  
 RP VARIANT OI-II CYS-926.  
 RX MEDLINE=88033031; PubMed=3667599;  
 RA Vogel B.E., Minor R.R., Freund M., Prockop D.J.;  
 RT "A point mutation in a type I procollagen gene converts glycine 748  
 of the alpha 1 chain to cysteine and destabilizes the triple helix in



DR InterPro; IPR001002; Chitin\_binding\_1.  
DR Pfam; PF00187; Chitin\_bind\_1; 4.  
DR PRINTS; PR00451; CHITINBINDING.  
DR ProDom; PD000609; Chitin\_binding\_1; 4.  
DR SMART; SM00270; ChtBD1; 4.  
DR PROSITE; PS00026; CHITIN\_BINDING; 4.  
KW Lectin; Repeat; Chitin-binding; 3D-structure; Signal.  
FT SIGNAL 1 26  
FT CHAIN 27 197 AGGLUTININ ISOLECTIN 1.  
FT PROPEP 198 212  
FT DOMAIN 27 69 CHITIN-BINDING 1.  
FT DOMAIN 70 112 CHITIN-BINDING 2.  
FT DOMAIN 113 155 CHITIN-BINDING 3.  
FT DOMAIN 156 197 CHITIN-BINDING 4.  
FT DISULFID 29 44  
FT DISULFID 38 50  
FT DISULFID 43 57  
FT DISULFID 61 66  
FT DISULFID 72 87  
FT DISULFID 81 93  
FT DISULFID 86 100  
FT DISULFID 104 109  
FT DISULFID 115 130  
FT DISULFID 124 136  
FT DISULFID 129 143  
FT DISULFID 147 152  
FT DISULFID 158 173  
FT DISULFID 167 179  
FT DISULFID 172 186  
FT DISULFID 190 195  
FT CONFLICT 63 63  
FT STRAND 30 30  
FT HELIX 31 33  
FT TURN 34 35  
FT STRAND 36 37  
FT HELIX 39 41  
FT STRAND 43 44  
FT TURN 46 47  
FT STRAND 50 52  
FT HELIX 54 57  
FT TURN 59 60  
FT STRAND 71 71  
FT HELIX 74 76  
FT TURN 77 78  
FT HELIX 82 84  
FT STRAND 86 87  
FT TURN 89 90  
FT STRAND 92 95  
FT HELIX 97 100  
FT TURN 102 103  
FT STRAND 106 106  
FT STRAND 114 114  
FT STRAND 116 116  
FT HELIX 117 119  
FT TURN 120 121  
FT STRAND 122 122  
FT HELIX 125 127  
FT STRAND 129 130  
FT TURN 132 133  
FT STRAND 135 138  
FT HELIX 140 143  
FT TURN 145 146  
FT STRAND 149 149  
FT STRAND 157 157  
FT STRAND 159 159  
FT HELIX 160 162  
FT TURN 163 164  
FT STRAND 165 165  
FT HELIX 168 170  
FT STRAND 172 173  
FT TURN 175 176  
FT STRAND 178 181  
FT HELIX 183 186

N -> D (IN REF. 2 AND 3).

FT TURN 188 189  
SQ SEQUENCE 212 AA; 21239 MW; EC7B6F007DDC15EB CRC64;  
Query Match 80.6%; Score 25; DB 1; Length 212;  
Best Local Similarity 83.3%; Pred. No. 24;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LMSTRA 6  
Db 3 MMSTRA 8  
Search completed: April 22, 2003, 12:51:47  
Job time : 5.5 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:24 ; Search time 6.875 Seconds  
(without alignments)  
97.882 Million cell updates/sec

Title: US-09-674-716B-5  
Perfect score: 31  
Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73: \*  
1: Pir1: \*  
2: Pir2: \*  
3: Pir3: \*  
4: Pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	74	2 G30538	Ig kappa chain V r
2	31	100.0	112	1 KVM516	Ig kappa chain V r
3	31	100.0	113	1 KVM551	Ig kappa chain V r
4	31	100.0	120	1 KVM567	Ig kappa chain pre
5	27	87.1	104	2 B69832	conserved hypother
6	27	87.1	234	2 S50468	hypothetical prote
7	27	87.1	342	2 C81451	flagellar motor sw
8	27	87.1	409	2 E86017	probable beta-keto
9	27	87.1	409	2 E91171	probable beta-keto
10	27	87.1	760	2 T39991	minichromosome mai
11	27	87.1	842	2 T16198	hypothetical prote
12	27	87.1	845	1 S34027	replication licens
13	27	87.1	1401	2 T39925	MAP kinase kinase
14	26	83.9	112	2 E95905	hypothetical prote
15	26	83.9	313	2 D96028	probable transcrip
16	26	83.9	341	1 QXFF2Y	NADH2 dehydrogenas
17	26	83.9	810	2 C70791	probable pona' pro
18	26	83.9	1453	2 S21626	collagen alpha 1(I
19	26	83.9	1464	1 CGHULS	collagen alpha 1(I
20	25	80.6	177	2 T51460	hypothetical prote
21	25	80.6	212	2 S09623	agglutinin isolect
22	25	80.6	212	2 T05936	agglutinin isolect
23	25	80.6	326	2 T26647	hypothetical prote
24	25	80.6	341	2 T39105	WD repeat protein
25	25	80.6	390	2 E98925	protein T22H9.4 [i
26	25	80.6	341	2 C56849	dopamine receptor-
27	25	80.6	511	2 AF0453	malate synthase (E
28	25	80.6	565	2 B72660	probable type II D
29	25	80.6	627	2 F95867	conserved hypother

alcohol oxidase (E  
probable membrane  
probable ATP-depen  
rifamycin polyketi  
gene MHC DQ-beta 1  
hypothetical prote  
conserved hypother  
hypothetical prote  
hypothetical prote  
IBIP8-4 protein [i  
hypothetical prote  
cytolysin II oporo  
emopamil-binding p  
periplasmic immuno  
conserved hypother  
Cof family protein

30 25 80.6 664 1 OXHOAP  
31 25 80.6 778 2 S56293  
32 25 80.6 887 2 T38885  
33 25 80.6 5069 2 T17464  
34 24 77.4 82 2 I36903  
35 24 77.4 144 2 G90913  
36 24 77.4 152 2 T44544  
37 24 77.4 164 2 T01726  
38 24 77.4 171 2 H96833  
39 24 77.4 171 2 G71422  
40 24 77.4 176 2 T10331  
41 24 77.4 190 4 A43860  
42 24 77.4 229 2 A56122  
43 24 77.4 250 2 AB3319  
44 24 77.4 270 2 G97904  
45 24 77.4 270 2 G95033

ALIGNMENTS

RESULT 1  
G30538  
Ig kappa chain V region (253.15E2) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Feb-1989 #sequence\_revision 10-Feb-1989 #text\_change 16-Aug-1996  
C:Accession: G30538  
R:Clafalin, J.L.; Berry, J.  
J. Immunol. 141, 4012-4019, 1988  
A:Title: Genetics of the phosphocholine-specific antibody response to Streptococcus pneu  
A:Reference number: A30534; MUID:89035545; PMID:3141511  
A:Accession: G30538  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
A:Molecule type: mRNA  
A:Residues: 1-74 <CIA>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 31; DB 2; Length 74;  
Best Local Similarity 100.0%; Pred. No. 0.75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q% 1 LMSTRAS 7  
Db 24 LMSTRAS 30

RESULT 2  
KVM516  
Ig kappa chain V region (M167) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Nov-1980 #sequence\_revision 30-Nov-1980 #text\_change 16-Aug-1996  
C:Accession: A01908  
R:Rudikoff, S.; Potter, M.  
Biochemistry 17, 2703-2707, 1978  
A:Title: kappa-Chain variable region from M167, a phosphorylcholine binding myeloma pro:  
A:Reference number: A01908; MUID:79000273; PMID:99160  
A:Accession: A01908  
A:Molecule type: protein  
A:Residues: 1-112 <RUD>  
C:Comment: This chain was isolated from a myeloma protein that binds phosphorylcholine.  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer  
F:16-95/Domain: immunoglobulin homology <IMM>  
F:23-93/Disulfide bonds: #status predicted

Query Match 100.0%; Score 31; DB 1; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q% 1 LMSTRAS 7

Db 55 LMSTRAS 61  
|||||

## RESULT 3

KVMS51  
Ig kappa chain V region (M511) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 21-Jan-2000  
C;Accession: A01910

R;Appella, E.

Mol. Immunol. 17, 711-718, 1980

A;Title: Amino acid sequence of the light chain variable region of M511, a phosphorylated  
A;Reference number: A01910; MUID:81052016; PMID:6776396

A;Accession: A01910

A;Molecule type: protein

A;Residues: 1-113 <APP>

C;Comment: This chain was isolated from a myeloma protein that binds phosphorylcholine.

C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer

F;16-95/Domain: immunoglobulin homology <IMM>

F;23-93/Disulfide bonds: #status predicted

Query Match 100.0%; Score 37; DB 1; Length 113;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

|||||

Db 55 LMSTRAS 61

## RESULT 4

KVMS67

Ig kappa chain precursor V region (VK167) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 21-Jan-2000  
C;Accession: A01909

R;Seising, E.; Storb, U.

Cell 25, 47-58, 1981

A;Title: Somatic mutation of immunoglobulin light-chain variable-region genes.

A;Reference number: A01909; MUID:82002223; PMID:6791832

A;Accession: A01909

A;Molecule type: DNA

A;Residues: 1-120 <SEL>

A;Note: the sequence was determined from the germline gene

C;Genetics:

A;Introns: 17/1

C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-120/Product: Ig kappa chain V region (VK167) #status predicted <MAT>

F;36-115/Domain: immunoglobulin homology <IMM>

F;43-113/Disulfide bonds: #status predicted

Query Match 100.0%; Score 31; DB 1; Length 120;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

|||||

Db 75 LMSTRAS 81

## RESULT 5

E69832

conserved hypothetical protein yhgB - Bacillus subtilis

N;Alternate names: hypothetical protein Y (pbpf 5' region)

C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
C;Accession: E69832; C40614

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galler

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Pardo, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Togato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: E69832

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-104 <KUN>

A;Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12849.1; PID:el183011,

A;Experimental source: strain 168

R;Popham, D.L.; Setlow, P.

J. Bacteriol. 175, 4870-4876, 1993

A;Title: Cloning, nucleotide sequence, and regulation of the Bacillus subtilis pbpf gene

A;Reference number: A40614; MUID:93328693; PMID:8335642

A;Accession: C40614

A;Molecule type: DNA

A;Residues: 1-67 <POP>

A;Cross-references: GB:LI0630

C;Genetics:

A;Gene: yhgB

Query Match 87.1%; Score 27; DB 2; Length 104;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

|||||

Db 38 LMSTRCS 44

## RESULT 6

S50468

hypothetical protein YER010c - yeast (Saccharomyces cerevisiae)

C;Species: Saccharomyces cerevisiae

C;Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 19-Apr-2002

C;Accession: S50468

R;Dietrich, F.S.

submitted to the EMBL Data Library, December 1994

A;Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and lambda

A;Reference number: S50433

A;Accession: S50468

A;Molecule type: DNA

A;Residues: 1-234 <DIE>

A;Cross-references: EMBL:U18778; NID:G603592; PID:G603602; GSPDB:GN000005; MIPS:YER010c

C;Genetics:

A;Gene: MIPS:YER010c

A;Cross-references: SGD:S0000812

A;Map position: 5R

Query Match 87.1%; Score 27; DB 2; Length 234;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 6

|||||

Db 102 LMSTRAS 107

## RESULT 7





A;Cross-references: EMBL:U42834; NID:g1125756; PID:g1125758; PIDN:AAA83583.1; CESP:F28B4  
 C;Genetics:  
 A;Gene: CESP:F28B4.2  
 A;Introns: 17/3; 59/3; 93/2; 116/2; 229/3; 333/2; 510/2; 659/3; 738/3; 803/2  
 C;Superfamily: CDC25-type guanine nucleotide exchange activator homology  
 F;212-481/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>  
 Query Match 87.1%; Score 27; DB 2; Length 842;  
 Best Local Similarity 100.0%; Pred. No. 94;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 MSTRAS 7  
 |||||  
 Db 650 MSTRAS 655  
 RESULT 12  
 S34027  
 replication licensing factor MCM7 [validated] - yeast (Saccharomyces cerevisiae)  
 N;Alternate names: cell division control protein CDC47; protein YBR1441; protein YBR202w  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 21-Jul-2000  
 C;Accession: S34027; S46074; S34925; S56049  
 R;Jacquet, M.  
 submitted to the EMBL Data Library, January 1993  
 A;Reference number: S34022  
 A;Accession: S34027  
 A;Molecule type: DNA  
 A;Residues: 1-845 <JAC>  
 A;Cross-references: EMBL:Z21487; NID:g311665; PIDN:CAA79689.1; PID:g311678  
 R;Bussereau, F.; Demolis, N.; Jacquet, M.; Walleet, L.  
 submitted to the Protein Sequence Database, August 1994  
 A;Reference number: S46054  
 A;Accession: S46074  
 A;Molecule type: DNA  
 A;Residues: 1-845 <BUS>  
 A;Cross-references: EMBL:Z36071; NID:g536576; PIDN:CAA85166.1; PID:g536577; GSPDB:GN0000  
 R;Bussereau, F.; Mallet, L.; Gaillon, L.; Jacquet, M.  
 Yeast 9, 797-806, 1993  
 A;Title: Yeast Sequencing Reports. A 12.8 kb segment, on the right arm of chromosome II  
 A;Reference number: S34925; MUID:93377417; PMID:8368014  
 A;Accession: S34925  
 A;Molecule type: DNA  
 A;Residues: 407-620 <BU2>  
 A;Cross-references: EMBL:Z21487  
 R;Dalton, S.  
 submitted to the EMBL Data Library, September 1994  
 A;Description: Cdc47 and Cdc54 belong to a family of proteins essential for initiation o  
 A;Reference number: S56049  
 A;Accession: S56049  
 A;Molecule type: DNA  
 A;Residues: 1-551, 'G', 553-555, 'TLN', 559-573, 'Y', 575-845 <DAL>  
 A;Cross-references: EMBL:U14730; NID:g508168; PIDN:AAA86309.1; PID:g508169  
 C;Comment: The complex of six MCM proteins is one of several proteins that must be bound  
 phosphorylated and dissociate from the chromatin.  
 C;Genetics:  
 A;Gene: SGD:CDC47; MIPS:YBR202w  
 A;Cross-references: SGD:S0000406; MIPS:YBR202w  
 A;Map position: 2R  
 C;Complex: The predominant form is a heterohexameric of MCM2 (PIR:S45757), MCM3 (PIR:A3637  
 component of replication licensing factor.  
 C;Function:  
 A;Description: MCM7 is a component of the replication licensing factor that permits DNA  
 A;Superfamily: replication licensing factor MCM7; MCM homology  
 C;Keywords: cell cycle control; DNA replication initiation; heterohexameric; nucleus; phos  
 F;227-719/Domain: MCM homology <MCM>  
 Query Match 87.1%; Score 27; DB 1; Length 845;  
 Best Local Similarity 100.0%; Pred. No. 94;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 MSTRAS 7  
 |||||

Db 300 MSTRAS 305  
 RESULT 13  
 T39225  
 MAP kinase kinase kinase - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Jan-2000  
 C;Accession: T39225  
 R;Churcher, C.M.; Gentles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, August 1997  
 A;Reference number: Z21837  
 A;Accession: T39225  
 A;Status: Preliminary; translated from GB/EMBL/DBBJ  
 A;Molecule type: DNA  
 A;Residues: 1-1401 <CHU>  
 A;Cross-references: EMBL:Z98763; PIDN:CAB11500.1; GSPDB:GN00066; SPDB:SPAC9G1.02  
 A;Experimental source: strain 972h-; cosmid c9G1  
 C;Genetics:  
 A;Gene: SPDB:SPAC9G1.02  
 A;Map position: 1  
 C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 Query Match 87.1%; Score 27; DB 2; Length 1401;  
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LMSTRAS 7  
 |||||  
 Db 81 LMSTRAS 87  
 RESULT 14  
 E95905  
 hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymb  
 C;Species: Sinorhizobium meliloti  
 C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C;Accession: E95905  
 R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A;Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo  
 A;Reference number: A95842; MUID:21396508; PMID:11481431  
 A;Accession: E95905  
 A;Status: Preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-112 <KUR>  
 A;Cross-references: GB:AL591985; PIDN:CAC48909.1; PID:g15140382; GSPDB:GN00167  
 A;Experimental source: strain 1021, megaplasmid pSymb  
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler,  
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A;Reference number: A96039; MUID:21368234; PMID:11474104  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: Smb20530  
 A;Genome: plasmid  
 Query Match 83.9%; Score 26; DB 2; Length 112;  
 Best Local Similarity 71.4%; Pred. No. 22;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LMSTRAS 7  
 :|||:|  
 Db 5 IMSTRAS 11  
 RESULT 15  
 D96028  
 Probable transcription activator of the pca operon, LysR family protein [imported] - Sin

C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #ssequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C:Accession: D96028  
 R:Finan, T.M.; Weidner, S.; Wong, K.; Ruhmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo  
 A:Reference number: A95842; MUID:21396508; PMID:11481431  
 A:Accession: D96028  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-313 <KUR>  
 A:Cross-references: GB:AL591985; PIDN:CAC49892.1; PID:gl5141380; GSPDB:GN00167  
 A:Experimental source: strain 1021, megaplasmid pSymb  
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebault, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: pcaQ; SMb20580  
 A:Genome: plasmid

Query Match 83.9%; Score 26; DB 2; Length 313;  
 Best Local Similarity 85.7%; Pred. No. 63;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7  
 Db 199 LMPTRAS 205

Search completed: April 22, 2003, 12:54:47  
 Job time : 8.875 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:47:43 ; Search time 14 Seconds  
(without alignments)  
103.024 Million cell updates/sec

Title: US-09-674-716B-5  
Perfect score: 31  
Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	87.1	342	16 Q9PI10	Q9pi10 campylobact
2	27	87.1	409	16 Q8X5U5	Q8x5u5 escherichia
3	27	87.1	860	5 Q19852	Q19852 caenorhabdi
4	27	87.1	1024	11 Q9ESJ5	Q9esj5 mus musculu
5	27	87.1	1049	2 Q9XBP6	Q9xbp6 myxococcus
6	26	83.9	112	16 Q92W36	Q92w36 rhizobium m
7	26	83.9	284	4 Q16050	Q16050 homo sapien
8	26	83.9	287	6 Q95ND8	Q95nd8 bos taurus
9	26	83.9	313	16 Q92TL9	Q92tl9 rhizobium m
10	26	83.9	325	11 Q60785	Q60785 mus musculu
11	26	83.9	338	8 Q9MFO0	Q9mfg0 cochlicomyia
12	26	83.9	338	8 Q9B2J6	Q9b2j6 chrysomya c
13	26	83.9	340	8 Q9XMP1	Q9xmp1 ceratitis c
14	26	83.9	341	8 Q9MD59	Q9md59 drosophila
15	26	83.9	341	8 Q9MD83	Q9md83 drosophila
16	26	83.9	341	8 Q9MD72	Q9md72 drosophila

17	26	83.9	341	8 Q9MJC9	Q9mjc9 drosophila
18	26	83.9	341	8 Q9MGP1	Q9mgp1 drosophila
19	26	83.9	341	8 Q9MGN7	Q9mgm7 drosophila
20	26	83.9	341	8 Q9MGN2	Q9mgm2 drosophila
21	26	83.9	341	8 Q9MGL3	Q9mgl3 drosophila
22	26	83.9	481	13 Q90IJ1	Q90yj1 brachydanio
23	26	83.9	589	11 Q99LL6	Q99ll6 mus musculu
24	26	83.9	810	16 Q69650	Q69650 mycobacteri
25	26	83.9	1260	5 Q9GY27	Q9gy27 leishmania
26	26	83.9	1453	11 Q63079	Q63079 rattus norv
27	26	83.9	1461	4 Q76045	Q76045 homo sapien
28	25	80.6	135	2 Q8RQG7	Q8rqg7 bacillus ce
29	25	80.6	149	8 Q8WA98	Q8wa98 narceus ann
30	25	80.6	177	10 Q9LFF8	Q9lff8 arabidopsis
31	25	80.6	211	16 Q8VKC3	Q8vkc3 mycobacteri
32	25	80.6	263	7 Q9MX44	Q9mx44 ginglymosto
33	25	80.6	290	16 Q9F3E2	Q9f3e2 streptomyce
34	25	80.6	326	5 Q9XW66	Q9xw66 caenorhabdi
35	25	80.6	341	3 Q9UT39	Q9ut39 schizosacch
36	25	80.6	367	5 Q8SXA7	Q8sxa7 drosophila
37	25	80.6	383	5 Q9VNI7	Q9vni7 drosophila
38	25	80.6	390	5 Q9TXN8	Q9txn8 caenorhabdi
39	25	80.6	400	10 Q9LUQ7	Q9luq7 arabidopsis
40	25	80.6	508	12 Q9E234	Q9e234 helicoverpa
41	25	80.6	510	12 Q99GU9	Q99gu9 helicoverp
42	25	80.6	511	13 Q9PSA7	Q9psa7 fugu rubrip
43	25	80.6	532	16 Q8ZAR5	Q8zar5 yersinia pe
44	25	80.6	536	9 Q94MR7	Q94mr7 bacterioph
45	25	80.6	581	1 Q9P9C2	Q9p9c2 uncultured

#### ALIGNMENTS

RESULT 1

ID	Q9PI10	PRELIMINARY;	PRT;	342 AA.
AC	Q9PI10;			
DT	01-OCT-2000 (TRENBLrel. 15, Created)			
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)			
DT	01-JUN-2002 (TRENBLrel. 21, Last annotation update)			
DE	Flagellar motor switch protein.			
GN	FLIG OR CJO319.			
OS	Campylobacter jejuni.			
OC	Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;			
OC	Campylobacter.			
OX	NCBI_TaxID=197;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NCTC 11168;			
RX	MEDLINE=20150912; PubMed=10688204;			
RA	Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,			
RA	Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,			
RA	Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,			
RA	Quail M.A., Rajadream M.A., Rutherford K.M., van Vliet A.H.M.,			
RA	Whithead S., Barrall B.G.;			
RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni			
RT	reveals hypervariable sequences."			
RL	Nature 403:665-668(2000).			
DR	EMBL; AL139074; CAB72786.1; -.			
DR	HSSP; Q9WV63; 1QC7.			
DR	InterPro; IPR000090; Flg_Motor_Flg.			
DR	Pfam; PF01706; Flg-C; 1.			
DR	PRINTS; PRO0954; FLGMOTORFLIG.			
DR	TIGRFAME; TIGR00207; flig; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 342 AA; 38429 MW; ED4D214417C1524C CRC64;			

Query Match 87.1%; Score 27; DB 16; Length 342;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MSTRAS 7

Db 289 MSTRAS 294  
|||||

```

RESULT 2
Q8X5U5 PRELIMINARY; PRT; 409 AA.
AC Q8X5U5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative beta-ketoacyl-ACP synthase (Putative 3-oxoacyl-(acyl-carrier-
DE protein) synthase II).
DE Z4866 OR EC54341.
GN Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
ON NCBI_TaxID=83334;
[1]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posifal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimallanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533 (2001).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RMD 0509952;
RX MEDLINE=1156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kuzokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22 (2001).
RN [2]
SEQUENCE FROM N.A.
RC EMBL; AE005571; AKG58601.1; -.
DR EMBL; AP002565; BAB37764.1; -.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt C; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 44148 MW; A4DAB0AF11ED883 CRC64;

Query Match 87.1%; Score 27; DB 16; Length 409;
Best Local Similarity 100.0%; Pred. NO. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MSTRAS 7
Db 77 MSTRAS 82
|||||

RESULT 3
Q19852 PRELIMINARY; PRT; 860 AA.
AC Q19852;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 96.1 kDa protein.
GN F28B4.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
ON NCBI_TaxID=6239;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Leimbach D.;
RT "The sequence of C. elegans cosmid F28B4.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U42834; AAA83583.2; -.
DR HSSP; O61193; 1RLP.
DR InterPro; IPR000651; RasGEFN.
DR InterPro; IPR001895; RasGEFN_CDC25.
DR InterPro; IPR000159; RA_domain.
DR Pfam; PF00788; RA; 1.
DR Pfam; PF00617; RasGEF; 1.
DR Pfam; PF00618; RasGEFN; 1.
DR SMART; SM00314; RA; 1.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEFN; 1.
DR PROSITE; PS00720; GDS_CDC25; 1.
KW Hypothetical protein.
SQ SEQUENCE 860 AA; 96122 MW; 16089524FBC65CA5 CRC64;

Query Match 87.1%; Score 27; DB 5; Length 860;
Best Local Similarity 100.0%; Pred. NO. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MSTRAS 7
Db 668 MSTRAS 673
|||||

RESULT 4
Q9ESJ5 PRELIMINARY; PRT; 1024 AA.
AC Q9ESJ5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GluR-delta2 phlic-protein.
GN GRID2IP OR DELPHILIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX Miyagi Y., Yamashita T., Okuda K., Mishina M., Kawamoto S.;
RT "Delphinin: A novel PDZ-containing protein associates with the GluR-
RT delta2 subunit.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099933; AAG31020.1; -.
DR HSSP; P29476; IQAV.
DR MGD; MGI:2176213; Grid2ip.
DR InterPro; IPR003104; PDZ.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF02181; FH2; 1.
DR Pfam; PF00595; PDZ; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00498; FH2; 1.

```

```

DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS50106; PDZ; 1.
SQ SEQUENCE 1024 AA; 112578 MW; E318AFDER02F846A CRC64;

Query Match      87.1%; Score 27; DB 11; Length 1024;
Best Local Similarity 85.7%; Pred.No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
   |:|||||
Db 53 LVSTRAS 59

RESULT 5
Q9XBP6
ID Q9XBP6 PRELIMINARY; PRT; 1049 AA.
AC Q9XBP6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Serine/threonine kinase PKN8.
GN PKN8.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZF1;
RA Inouye S., Jain R., Ueki T., Nariya H., Xu C., Hsu M.,
RA Munoz-Dorado J., Farez-Vidal E., Inouye M.;
RT "Sequence Analysis of 13 Eukaryotic-like Protein Ser/Thr Kinases of
RT Myxococcus xanthus, a Developmental Bacterium and Significance of
RT Their Coexistence with Protein His Kinases";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR DR EMBL; AF159691; AAD42856.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00515; TPR; 5.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 1049 AA; 114313 MW; 7752862DAA25338C CRC64;

Query Match      87.1%; Score 27; DB 2; Length 1049;
Best Local Similarity 85.7%; Pred.No. 2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
   |:|||||
Db 1002 LMEIRAS 1008

RESULT 6
Q92W36
ID Q92W36 PRELIMINARY; PRT; 112 AA.
AC Q92W36;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-WAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein RB0509.
GN RB0509 OR SMB20530.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;

```

```

KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 313 AA; 33832 MW; 4787BC1BE5476709 CRC64;

Query Match      83.9%; Score 26; DB 16; Length 313;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
   |||||
DB 199 LMPTRAS 205

RESULT 10
Q60785 PRELIMINARY; PRT; 325 AA.
ID O60785;
AC O60785;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Alpha-1 type I procollagen (Fragment).
GN COL1A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Z., Laptiev A.V., Prockop D.J.;
RT "The nucleotide sequence of cDNA for the last 5 exons and 3' -
RL untranslated region of murine type I(1) procollagen.";
DR EMBL; U03419; AAA03475.1; -
DR MGD; MGI:88467; Colla1.
DR InterPro; IPR00087; Collagen.
DR InterPro; IPR00085; Fib_collagen_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 1.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
KW Collagen.
FT NON TER
FT NON TER
SQ SEQUENCE 325 AA; 35229 MW; A5A21E74DFDE3BF1 CRC64;

Query Match      83.9%; Score 26; DB 11; Length 325;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
   |||||
DB 218 LMSTEAS 224

RESULT 11
Q9MFQ0 PRELIMINARY; PRT; 338 AA.
ID Q9MFQ0;
AC Q9MFQ0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE NADH subunit 2.
OS Cochliomyia hominivorax (Primary screw-worm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eurygaster; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Cecidoidea; Calliphoridae; Cochliomyia.
OC NCBI_TaxID=115425;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20485491; PubMed=11029671;
RA Lessinger A.C., Martins Junqueira A.C., Lemos T.A., Kemper E.L.,
RA da Silva F.R., Vettore A.L., Arruda P., Azeredo-Espin A.M.;
RT "The mitochondrial genome of the primary screwworm fly Cochliomyia
RT hominivorax (Diptera: Calliphoridae).";

KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 313 AA; 33832 MW; 4787BC1BE5476709 CRC64;

Query Match      83.9%; Score 26; DB 16; Length 313;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
   |||||
DB 199 LMPTRAS 205

RESULT 10
Q60785 PRELIMINARY; PRT; 325 AA.
ID O60785;
AC O60785;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Alpha-1 type I procollagen (Fragment).
GN COL1A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Z., Laptiev A.V., Prockop D.J.;
RT "The nucleotide sequence of cDNA for the last 5 exons and 3' -
RL untranslated region of murine type I(1) procollagen.";
DR EMBL; U03419; AAA03475.1; -
DR MGD; MGI:88467; Colla1.
DR InterPro; IPR00087; Collagen.
DR InterPro; IPR00085; Fib_collagen_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 1.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
KW Collagen.
FT NON TER
FT NON TER
SQ SEQUENCE 325 AA; 35229 MW; A5A21E74DFDE3BF1 CRC64;

Query Match      83.9%; Score 26; DB 11; Length 325;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
   |||||
DB 218 LMSTEAS 224

RESULT 11
Q9MFQ0 PRELIMINARY; PRT; 338 AA.
ID Q9MFQ0;
AC Q9MFQ0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE NADH subunit 2.
OS Cochliomyia hominivorax (Primary screw-worm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eurygaster; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Cecidoidea; Calliphoridae; Cochliomyia.
OC NCBI_TaxID=115425;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20485491; PubMed=11029671;
RA Lessinger A.C., Martins Junqueira A.C., Lemos T.A., Kemper E.L.,
RA da Silva F.R., Vettore A.L., Arruda P., Azeredo-Espin A.M.;
RT "The mitochondrial genome of the primary screwworm fly Cochliomyia
RT hominivorax (Diptera: Calliphoridae).";

```



```

RL Insect Mol. Biol. 9:521-529(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Lessinger A.C., Junqueira A.C.M., Lemos T.A., Kemper E.L.,
RA Vettore A.L., da Silva F.R., Arruda P., Azeredo-Espin A.M.L.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL; AF260826; AAF78613.1; -.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
KW SEQUENCE 338 AA; 39337 MW; E6501A86D02439D0 CRC64;
SQ

Query Match 83.9%; Score 26; DB 8; Length 338;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 50 LMSTEAS 56

RESULT 12
Q9B2J6 PRELIMINARY; PRT; 338 AA.
AC Q9B2J6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE NADH subunit 2.
GN ND2.
OS Chrysomya chloropyga.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Cecidoidea; Calliphoridae; Chrysomya.
OX NCBI_TaxID=142378;
RN [1]
RP SEQUENCE FROM N.A.
RA Junqueira A.C.M.;
RL "The complete mitochondrial genome of the Chrysomya putoria.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL; AF352790; AAK21318.1; -.
DR InterPro; IPR003917; NADHUB_oxred2.
DR Pfam; PF00361; oxidored_q1.
DR PRINTS; PR01436; NADHDHGNASE2.
DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
KW SEQUENCE 338 AA; 39228 MW; ABD5B1B5A5A46A23 CRC64;
SQ

Query Match 83.9%; Score 26; DB 8; Length 338;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 50 LMSTEAS 56

RESULT 13
Q9XMP1 PRELIMINARY; PRT; 340 AA.
AC Q9XMP1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE NADH dehydrogenase (ubiquinone) chain 2 (EC 1.6.5.3).
GN NDH-U1.
OS Ceratitis capitata (Mediterranean fruit fly).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```

---

```

OC Tephritoidea; Tephritidae; Ceratitis.
OX NCBI_TaxID=7213;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20225575; PubMed=10762421;
RA Spanos L., Koutroumbas G., Kotsyfakis M., Louis C.;
RT "The complete sequence of the mitochondrial genome of the Medfly,
RT Ceratitis capitata."
RL Insect Mol. Biol. 9:139-144(2000).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL; AJ242872; CAB45088.1; -.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
KW SEQUENCE 340 AA; 39231 MW; F4C646E1C5E3F4F CRC64;
SQ

Query Match 83.9%; Score 26; DB 8; Length 340;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 50 LMSTEAS 56

RESULT 14
Q9MDS9 PRELIMINARY; PRT; 341 AA.
AC Q9MDS9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE NADH dehydrogenase subunit 2.
DE Drosophila simulans (Fruit fly).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7240;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=RU35, C167, DSW, DSR, MD106, MD225, AND SC00;
RA Ballard J.W.O.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL; AF200845; AAF77447.1; -.
DR EMBL; AF200839; AAF77369.1; -.
DR EMBL; AF200840; AAF77382.1; -.
DR EMBL; AF200841; AAF77395.1; -.
DR EMBL; AF200842; AAF77408.1; -.
DR EMBL; AF200843; AAF77421.1; -.
DR EMBL; AF200844; AAF77434.1; -.
DR FlyBase; FBgn0012880; Deim\mt.ND2.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
KW SEQUENCE 341 AA; 39734 MW; 8C42C96FF977A173 CRC64;
SQ

Query Match 83.9%; Score 26; DB 8; Length 341;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 51 LMSTEAS 57

RESULT 15
Q9MD83 PRELIMINARY; PRT; 341 AA.
AC Q9MD83;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE NADH dehydrogenase subunit 2.  
 OS Drosophila simulans (Fruit fly), and  
 OS Drosophila mauritiana (Fruit fly).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7240, 7226;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC SPECIES=D.simulans, and D.mauritiana; STRAIN=VARIOUS STRAINS;  
 RA Ballard J.W.O.;

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases  
 CC -|- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.

DR EMBL; AF200854; AAF77564.1; -;  
 DR EMBL; AF200831; AAF77265.1; -;  
 DR EMBL; AF200846; AAF77460.1; -;  
 DR EMBL; AF200847; AAF77473.1; -;  
 DR EMBL; AF200848; AAF77486.1; -;  
 DR EMBL; AF200849; AAF77499.1; -;  
 DR EMBL; AF200850; AAF77512.1; -;  
 DR EMBL; AF200851; AAF77525.1; -;  
 DR EMBL; AF200852; AAF77538.1; -;  
 DR EMBL; AF200853; AAF77551.1; -;  
 DR FlyBase; FBgn0012512; Dmau\mt:ND2.  
 DR FlyBase; FBgn0012880; Dsim\mt:ND2.  
 DR InterPro; IPR001750; Oxidored\_q1.  
 DR Pfam; PF00361; Oxidored\_q1; 1\_  
 DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.  
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.  
 SQ SEQUENCE 341 AA, 39649 MW, 27BAE96F482EC088 CRC64;

Query Match 83.9%; Score 26; DB 8; Length 341;  
 Best Local Similarity 85.7%; Pred. NO. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
 Db 51 LMSTEAS 57

Search completed: April 22, 2003, 12:53:48  
 Job time : 17 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:45:19 ; Search time 17.75 Seconds  
(without alignments)  
52.550 Million cell updates/sec

Title: US-09-674-716B-5

Perfect score: 31

Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq 101002:\*

- 1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	7	19	AAW39819
2	31	100.0	7	19	AAW39822
3	31	100.0	7	19	AAW39825
4	31	100.0	7	21	AAW32255
5	31	100.0	100	22	AAE06969
6	31	100.0	113	19	AAW39886
7	31	100.0	113	19	AAW39882
8	31	100.0	113	19	AAW39802
9	31	100.0	113	19	AAW39803
10	31	100.0	113	19	AAW39804

11	11	100.0	116	21	AAW32262	Humanised anti-CD2
12	31	100.0	145	21	AAW32261	Mouse anti-CD23 MA
13	28	90.3	7	19	AAW39816	Light chain CDR2 o
14	28	90.3	113	19	AAW39801	Variable domain of
15	28	90.3	274	19	AAW39899	Single chain Fv re
16	27	87.1	7	19	AAW39876	Light chain CDR2 o
17	27	87.1	68	22	AAU64213	Protonibacterium
18	27	87.1	123	22	AAU54742	Protonibacterium
19	27	87.1	131	12	AAU12332	Mouse Mab 2E12 L c
20	27	87.1	132	12	AAU12354	Light (kappa) chai
21	27	87.1	140	22	AAU95756	Human reproductive
22	27	87.1	409	22	ABB52717	Escherichia coli p
23	26	83.9	77	22	AAU39838	Protonibacterium
24	26	83.9	115	22	ABB27906	Human peptide #557
25	26	83.9	115	22	ABB33078	Peptide #584 encod
26	26	83.9	115	22	ABB18546	Protein #545 encod
27	26	83.9	115	22	AAU53875	Human brain expres
28	26	83.9	115	22	AAU66263	Human bone marrow
29	26	83.9	115	22	AAU14133	Peptide #567 encod
30	26	83.9	115	22	AAU26539	Peptide #576 encod
31	26	83.9	115	22	AAU01871	Peptide #553 encod
32	26	83.9	115	23	ABG35910	Human peptide enco
33	26	83.9	134	22	ABG10627	Novel human diagno
34	26	83.9	250	18	AAU12845	Pro-alpha(I) chai
35	26	83.9	441	22	AAU75593	Human colon cancer
36	26	83.9	449	21	AAU43439	Human cancer assoc
37	26	83.9	810	22	AAU81232	Mycobacterium tube
38	26	83.9	980	22	ABG03181	Novel human diagno
39	26	83.9	980	22	ABG10626	Novel human diagno
40	26	83.9	1341	16	AAU71701	Collagen alpha 1 (
41	26	83.9	1341	21	AAU96122	Collagen type I al
42	26	83.9	1341	23	ABB80733	Collagen type I-al
43	26	83.9	1341	23	ABB09625	Amino acid sequenc
44	26	83.9	1341	23	ABB16475	Human collagen alp
45	26	83.9	1411	21	AAU56800	Human preproalpha

#### ALIGNMENTS

RESULT 1  
AAW39819  
ID AAW39819 standard; peptide; 7 AA.  
XX AC AAW39819;  
XX DT 16-JUN-1998 (first entry)  
XX DE Light chain CDR2 of catalytic antibody 6A12.  
XX DE Variable domain; lambda light chain; catalytic antibody; degradation;  
XX KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
XX KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
XX KW overdose; addiction.  
XX OS Mus sp.  
XX PN WO9749800-A1.  
XX PD 31-DEC-1997.  
XX PF 25-JUN-1997; 97WO-US10965.  
XX PR 25-JUN-1996; 96US-0672345.  
XX PA (UYCO ) UNIV COLUMBIA NEW YORK.  
XX PI Landry DW;  
XX DR WPI; 1998-077166/07.  
XX PT New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required

PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding

PS Claim 13; Page 81; 147pp; English.

XX AAW39818-20 represent the sequences of the light chain complementarity  
CC determining regions (CDRs) of the catalytic antibody 6A12, which is able  
CC to degrade cocaine. A series of cocaine transition state analogues  
CC (TSAs) were prepared and used to immunise mice for production of  
CC hybridomas. Catalytic antibodies were identified by their capacity to  
CC release 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody was  
CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
CC monoester transition state analogue. Antibody 6A12 has a per minute Kcat  
CC of 0.072. The antibodies reduce the concentration of cocaine in a  
CC subject, and are used particularly for the treatment of cocaine in an  
CC are also used for treating addiction (by reducing the in vivo  
CC concentration that can be achieved).

SQ Sequence 7 AA;

Query Match 100.0%; Score 31; DB 19; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

|||||

Db 1 LMSTRAS 7

RESULT 2

AAW39822

ID AAW39822 standard; peptide; 7 AA.

AC AAW39822;

XX 16-JUN-1998 (first entry)

DT Light chain CDR2 of catalytic antibody 2A10.

DE

XX Variable domain; lambda light chain; catalytic antibody; degradation;

KW cocaine; cocaine transition state analogue; TSA; benzoic acid;

KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;

KW overdose; addiction.

XX Mus sp.

OS WO9749800-A1.

XX 31-DEC-1997.

XX 25-JUN-1997; 97WO-US10965.

XX 25-JUN-1996; 96US-0672345.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Landry DW;

XX WPI; 1998-077166/07.

XX New catalytic antibodies able to decompose cocaine, single-chain

PT analogues - used to treat cocaine overdose and addiction, required

PT in far smaller doses than antibodies that antagonise cocaine by

PT simply binding

XX Claim 15; Page 82; 147pp; English.

XX AAW39821-23 represent the sequences of the light chain complementarity

CC determining regions (CDRs) of the catalytic antibody 2A10, which is able

CC to degrade cocaine. A series of cocaine transition state analogues

CC (TSAs) were prepared and used to immunise mice for production of

CC hybridomas. Catalytic antibodies were identified by their capacity to

CC release 3H-benzoic acid from 3H-phenyl cocaine. The 2A10 antibody was

CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
CC monoester transition state analogue. Antibody 2A10 has a per minute Kcat  
CC of 0.011. The antibodies reduce the concentration of cocaine in a  
CC subject, and are used particularly for the treatment of cocaine in an  
CC are also used for treating addiction (by reducing the in vivo  
CC concentration that can be achieved).

XX Sequence 7 AA;

Query Match 100.0%; Score 31; DB 19; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

|||||

Db 1 LMSTRAS 7

RESULT 3

AAW39825

ID AAW39825 standard; peptide; 7 AA.

XX AAW39825;

XX 16-JUN-1998 (first entry)

DT Light chain CDR2 of catalytic antibody 12H1.

DE

XX Variable domain; lambda light chain; catalytic antibody; degradation;

KW cocaine; cocaine transition state analogue; TSA; benzoic acid;

KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;

KW overdose; addiction.

XX Mus sp.

OS WO9749800-A1.

XX 31-DEC-1997.

XX 25-JUN-1997; 97WO-US10965.

XX 25-JUN-1996; 96US-0672345.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Landry DW;

XX WPI; 1998-077166/07.

XX New catalytic antibodies able to decompose cocaine, single-chain

PT analogues - used to treat cocaine overdose and addiction, required

PT in far smaller doses than antibodies that antagonise cocaine by

PT simply binding

XX Claim 17; Page 83; 147pp; English.

XX AAW39824-26 represent the sequences of the light chain complementarity

CC determining regions (CDRs) of the catalytic antibody 12H1, which is able

CC to degrade cocaine. A series of cocaine transition state analogues

CC (TSAs) were prepared and used to immunise mice for production of

CC hybridomas. Catalytic antibodies were identified by their capacity to

CC release 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody was

CC identified using TSA2, and has a per minute Kcat of 0.16. The antibodies

CC reduce the concentration of cocaine in a subject, and are used

CC particularly for the treatment of cocaine in an overdose. They are also used for

CC treating addiction (by reducing the in vivo concentration that can be

CC achieved).

XX Sequence 7 AA;

Query Match 100.0%; Score 31; DB 19; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 31; DB 21; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels

The patent discloses a humanised antibody or its antigen-binding fragment, having binding specificity for CC-chemokine receptor 2 (CCR2), comprising an antigen binding region of non-human origin (CCR2), and at least a portion of an immunoglobulin of human origin. The humanised antibodies are useful for inhibiting the interaction of a cell expressing CCR2. They are useful for inhibiting or treating HIV infection. The proteins of the invention are useful for inhibiting leukocyte trafficking, for treating CCR2-mediated disorders such as inflammatory disorder, autoimmune disorders such as rheumatoid arthritis and multiple sclerosis, atherogenesis and atherosclerosis, and for inhibiting restenosis. They are useful in therapy or diagnosis, and in the manufacture of a medicament for treating CCR-2 mediated disease. They are also useful for treating allergy, anaphylaxis, malignancy, chronic and acute inflammation, histamine and IgE-mediated allergic reaction, shock, stenosis, allograft rejection, fibrotic disease, asthma, inflammatory glomerulopathies, acquired immune deficiency syndrome (AIDS), restenosis associated with vascular intervention, including angioplasty and/or stent placement in a mammal. Humanised antibodies are also useful for inhibiting narrowing of the lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of a vessel in a mammal, preferably associated with vascular intervention. The present sequence is mouse germline kappa light chain variable (VK) region, 167/24.

35X



OS Mus sp.  
XX WO9749800-A1.  
XX 31-DEC-1997.  
XX 25-JUN-1997; 97WO-US10965.  
XX 25-JUN-1996; 96US-0672345.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Landry DW;  
XX WPI; 1998-077166/07.  
XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding  
XX Claim 14; Page 72; 147pp; English.  
XX AA39801-05 represent the amino acid sequences of the variable domain  
CC of the kappa light chain of catalytic antibodies which are able to  
CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody (AA39807  
CC represents the heavy chain) was identified using TSA1, which is an  
CC immunogenic conjugate of a phosphate monoester transition state  
CC analogue. Antibody 6A12 has a per minute Kcat of 0.072. The antibodies  
CC reduce the concentration of cocaine in a subject, and are used  
CC particularly for the treatment of an overdose. They are also used  
CC for treating addiction (by reducing the in vivo concentration that can  
CC be achieved).  
XX SQ Sequence 113 AA;  
Query Match 100.0%; Score 31; DB 19; Length 113;  
Best Local Similarity 100.0%; Pred. NO. 6.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 55 LMSTRAS 61  
RESULT 9  
AAW39803  
ID AA39803 standard; protein; 113 AA.  
XX AC AAW39803;  
XX 16-JUN-1998 (first entry)  
XX Variable domain of the kappa light chain of catalytic antibody 12H1.  
XX Variable domain; lambda light chain; catalytic antibody; degradation;  
XX cocaine; cocaine transition state analogue; TSA; benzoic acid;  
XX phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
XX overdose; addition.  
XX Mus sp.  
XX WO9749800-A1.  
XX 31-DEC-1997.  
XX 25-JUN-1997; 97WO-US10965.  
XX 25-JUN-1996; 96US-0672345.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Landry DW;  
XX WPI; 1998-077166/07.  
XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding  
XX Claim 14; Page 72; 147pp; English.  
XX AA39801-05 represent the amino acid sequences of the variable domain  
CC of the kappa light chain of catalytic antibodies which are able to  
CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody (AAW39807  
CC represents the heavy chain) was identified using TSA1, which is an  
CC immunogenic conjugate of a phosphate monoester transition state  
CC analogue. Antibody 6A12 has a per minute Kcat of 0.072. The antibodies  
CC reduce the concentration of cocaine in a subject, and are used  
CC particularly for the treatment of an overdose. They are also used  
CC for treating addiction (by reducing the in vivo concentration that can  
CC be achieved).  
XX SQ Sequence 113 AA;  
Query Match 100.0%; Score 31; DB 19; Length 113;  
Best Local Similarity 100.0%; Pred. NO. 6.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 55 LMSTRAS 61  
RESULT 9  
AAW39803  
ID AA39803 standard; protein; 113 AA.  
XX AC AAW39803;  
XX 16-JUN-1998 (first entry)  
XX Variable domain of the kappa light chain of catalytic antibody 12H1.  
XX Variable domain; lambda light chain; catalytic antibody; degradation;  
XX cocaine; cocaine transition state analogue; TSA; benzoic acid;  
XX phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
XX overdose; addition.  
XX Mus sp.  
XX WO9749800-A1.  
XX 31-DEC-1997.  
XX 25-JUN-1997; 97WO-US10965.  
XX 25-JUN-1996; 96US-0672345.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Landry DW;  
XX WPI; 1998-077166/07.  
XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding  
XX Claim 14; Page 72; 147pp; English.  
XX AA39801-05 represent the amino acid sequences of the variable domain  
CC of the kappa light chain of catalytic antibodies which are able to  
CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody (AAW39807  
CC represents the heavy chain) was identified using TSA1, which is an  
CC immunogenic conjugate of a phosphate monoester transition state  
CC analogue. Antibody 6A12 has a per minute Kcat of 0.072. The antibodies  
CC reduce the concentration of cocaine in a subject, and are used  
CC particularly for the treatment of an overdose. They are also used  
CC for treating addiction (by reducing the in vivo concentration that can  
CC be achieved).  
XX SQ Sequence 113 AA;  
Query Match 100.0%; Score 31; DB 19; Length 113;  
Best Local Similarity 100.0%; Pred. NO. 6.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 55 LMSTRAS 61  
RESULT 10  
AAW39804  
ID AAW39804 standard; protein; 113 AA.  
XX AC AAW39804;  
XX 16-JUN-1998 (first entry)  
XX Variable domain of the kappa light chain of catalytic antibody 2A10.  
XX Variable domain; lambda light chain; catalytic antibody; degradation;  
XX cocaine; cocaine transition state analogue; TSA; benzoic acid;  
XX phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
XX overdose; addition.  
XX Mus sp.  
XX WO9749800-A1.  
XX 31-DEC-1997.  
XX 25-JUN-1997; 97WO-US10965.  
XX 25-JUN-1996; 96US-0672345.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Landry DW;  
XX WPI; 1998-077166/07.  
XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding  
XX Claim 18; Page 73; 147pp; English.  
XX AA39801-05 represent the amino acid sequences of the variable domain  
CC of the kappa light chain of catalytic antibodies which are able to  
CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody (AAW39808  
CC represents the heavy chain) was identified using TSA2, and has a per  
CC minute Kcat of 0.016. The antibodies reduce the concentration of cocaine  
CC in a subject, and are used particularly for the treatment of an  
CC overdose. They are also used for treating addiction (by reducing the in  
CC vivo concentration that can be achieved).  
XX SQ Sequence 113 AA;  
Query Match 100.0%; Score 31; DB 19; Length 113;  
Best Local Similarity 100.0%; Pred. NO. 6.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 55 LMSTRAS 61

PA (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Landry DW;  
XX WPI; 1998-077166/07.  
XX P-PSDB; AAV09802.  
XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding  
XX Claim 18; Page 73; 147pp; English.  
XX AA39801-05 represent the amino acid sequences of the variable domain  
CC of the kappa light chain of catalytic antibodies which are able to  
CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody (AAW39808  
CC represents the heavy chain) was identified using TSA2, and has a per  
CC minute Kcat of 0.016. The antibodies reduce the concentration of cocaine  
CC in a subject, and are used particularly for the treatment of an  
CC overdose. They are also used for treating addiction (by reducing the in  
CC vivo concentration that can be achieved).  
XX SQ Sequence 113 AA;  
Query Match 100.0%; Score 31; DB 19; Length 113;  
Best Local Similarity 100.0%; Pred. NO. 6.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 55 LMSTRAS 61  
RESULT 10  
AAW39804  
ID AAW39804 standard; protein; 113 AA.  
XX AC AAW39804;  
XX 16-JUN-1998 (first entry)  
XX Variable domain of the kappa light chain of catalytic antibody 2A10.  
XX Variable domain; lambda light chain; catalytic antibody; degradation;  
XX cocaine; cocaine transition state analogue; TSA; benzoic acid;  
XX phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
XX overdose; addition.  
XX Mus sp.  
XX WO9749800-A1.  
XX 31-DEC-1997.  
XX 25-JUN-1997; 97WO-US10965.  
XX 25-JUN-1996; 96US-0672345.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Landry DW;  
XX WPI; 1998-077166/07.  
XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding  
XX Claim 18; Page 73; 147pp; English.  
XX AA39801-05 represent the amino acid sequences of the variable domain  
CC of the kappa light chain of catalytic antibodies which are able to  
CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody (AAW39808  
CC represents the heavy chain) was identified using TSA2, and has a per  
CC minute Kcat of 0.016. The antibodies reduce the concentration of cocaine  
CC in a subject, and are used particularly for the treatment of an  
CC overdose. They are also used for treating addiction (by reducing the in  
CC vivo concentration that can be achieved).  
XX SQ Sequence 113 AA;  
Query Match 100.0%; Score 31; DB 19; Length 113;  
Best Local Similarity 100.0%; Pred. NO. 6.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 55 LMSTRAS 61  
RESULT 10  
AAW39804  
ID AAW39804 standard; protein; 113 AA.  
XX AC AAW39804;  
XX 16-JUN-1998 (first entry)  
XX Variable domain of the kappa light chain of catalytic antibody 2A10.  
XX Variable domain; lambda light chain; catalytic antibody; degradation;  
XX cocaine; cocaine transition state analogue; TSA; benzoic acid;  
XX phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
XX overdose; addition.  
XX Mus sp.  
XX WO9749800-A1.  
XX 31-DEC-1997.  
XX 25-JUN-1997; 97WO-US10965.  
XX 25-JUN-1996; 96US-0672345.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Landry DW;  
XX WPI; 1998-077166/07.  
XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding  
XX Claim 18; Page 73; 147pp; English.  
XX AA39801-05 represent the amino acid sequences of the variable domain  
CC of the kappa light chain of catalytic antibodies which are able to  
CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody (AAW39808  
CC represents the heavy chain) was identified using TSA2, and has a per  
CC minute Kcat of 0.016. The antibodies reduce the concentration of cocaine  
CC in a subject, and are used particularly for the treatment of an  
CC overdose. They are also used for treating addiction (by reducing the in  
CC vivo concentration that can be achieved).  
XX SQ Sequence 113 AA;  
Query Match 100.0%; Score 31; DB 19; Length 113;  
Best Local Similarity 100.0%; Pred. NO. 6.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 55 LMSTRAS 61

PS Claim 16; Pages 73-74; 147pp; English.

XX AAW39801-05 represent the amino acid sequences of the variable domain  
 CC of the Kappa light chain of catalytic antibodies which are able to  
 CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
 CC were prepared and used to immunise mice for production of hybridomas.  
 CC Catalytic antibodies were identified by their capacity to release  
 CC 3H-benzoic acid from 3H-phenyl cocaine. The 2A10 antibody (AAW39809  
 CC represents the heavy chain) was identified using TSA1, which is an  
 CC immunogenic conjugate of a phosphate monoester transition state  
 CC analogue. Antibody 2A10 has a per minute Kcat of 0.011. The antibodies  
 CC reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used  
 CC for treating addiction (by reducing the in vivo concentration that can  
 CC be achieved).

XX SQ Sequence 113 AA;  
 Query Match 100.0%; Score 31; DB 19; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
 DB 55 LMSTRAS 61  
 |||||

RESULT 11

AAV32262  
 ID AAY32262 standard; Protein; 116 AA.

AC AAY32262;

XX 15-FEB-2000 (first entry)

XX Humanised anti-CD23 Mab C11 light chain variable region.

XX CD23; FCERII; IgE receptor; monoclonal antibody; C11; mouse; human;  
 KW monoclonal antibody; chimeric antibody; humanised antibody;  
 KW complementarity determining region; CDR; autoimmune disease;  
 KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
 KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
 KW urticaria; nephrotic syndrome; glomerulonephritis;  
 KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
 KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
 KW graft-versus-host disease; COPD; bronchitis; diabetes;  
 KW B-cell malignancy; therapy.

XX Homo sapiens.  
 OS Synthetic.

XX Key Location/Qualifiers  
 FT Region 1..23 "framework region 1"  
 FT Region 24..39 "framework region 1"  
 FT Region 40..54 "framework region 2"  
 FT Region 55..61 "framework region 2"  
 FT Region 62..93 "framework region 3"  
 FT Region 94..102 "framework region 3"  
 FT Region 103..113 "framework region 4"  
 FT Region 103..113 "framework region 4"

XX WO9558679-A1.

XX 18-NOV-1999.

XX 07-MAY-1999; 99WO-GB01434.

XX

PR 09-MAY-1998; 98GB-0009839.  
 XX (GLAX ) GLAXO GROUP LTD.  
 PA Bonney JWP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 PI N-PSDB; AAZ34747.  
 XX WPI; 2000-053101/04.  
 DR N-PSDB; AAZ34747.  
 XX Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 PT Claim 9; Fig 3; 8pp; English.

XX This sequence represents the light chain variable region (VL) of  
 CC humanised anti-CD23 (FCERII) monoclonal antibody C11, composed of  
 CC a human framework (HSIGKVII) and the light chain complementarity  
 CC determining regions (see AAY32254-56) of murine antibody C11. The  
 CC DNA was constructed by splice overlap PCR. The invention provides  
 CC altered antibodies, such as chimeric or humanised antibodies, which  
 CC comprise sufficient of the amino acid sequences of the C11 light  
 CC and heavy chain complementarity determining regions to render them  
 CC capable of binding to the CD23 type II molecule expressed on  
 CC haematopoietic cells. The antibodies are used to block soluble  
 CC CD23 formation in human therapy, for the treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis,  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.

XX SQ Sequence 116 AA;

Query Match 100.0%; Score 31; DB 21; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 6.8;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
 DB 55 LMSTRAS 61  
 |||||

RESULT 12

AAV32261  
 ID AAY32261 standard; Protein; 145 AA.

XX AAY32261;

XX 15-FEB-2000 (first entry)

XX Mouse anti-CD23 Mab C11 light chain variable region.

XX CD23; FCERII; IgE receptor; monoclonal antibody; C11; mouse;  
 KW monoclonal antibody; chimeric antibody; humanised antibody;  
 KW complementarity determining region; CDR; autoimmune disease;  
 KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
 KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
 KW urticaria; nephrotic syndrome; glomerulonephritis;  
 KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
 KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
 KW graft-versus-host disease; COPD; bronchitis; diabetes;  
 KW B-cell malignancy; therapy.

XX Mus musculus.

XX Key Location/Qualifiers  
 FT Region 55..70  
 FT /note= "CDR LI"



FT Region 83..92  
FT /note= "CDR L2"  
FT 125..134  
FT /note= "CDR L3"  
XX  
XX  
PN WO958679-A1.  
XX  
XX 18-NOV-1999.  
XX  
XX 07-MAY-1999; 99WO-GB01434.  
XX  
XX 09-MAY-1998; 98GB-0009839.  
XX  
XX (GLAX ) GLAXO GROUP LTD.  
XX  
XX Bonney J, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
XX WPI; 2000-053101/04.  
XX N-PSDB; AA234746.  
XX  
XX Cell receptor specific antibodies useful for treating e.g. arthritis,  
XX diabetes, multiple sclerosis and psoriasis -  
XX Claim 8; Fig 2; 81pp; English.  
XX  
XX This sequence represents the light chain variable region (VL) of  
XX murine anti-CD23 (FCER11) monoclonal antibody C11. The invention  
XX provides altered antibodies, such as chimeric or humanised  
XX antibodies (see AA32262 and AA32263), which comprise sufficient of  
XX the amino acid sequences of the C11 light and heavy chain  
XX complementarity determining regions (see AA32254-59) to render them  
XX capable of binding to the CD23 type II molecule expressed on  
XX haematopoietic cells. The antibodies are used to block soluble  
XX CD23 formation in human therapy, for the treatment of arthritis,  
XX lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
XX diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
XX syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
XX colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
XX asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
XX eczema, graft-versus-host disease, COPD, sinusitis, bronchitis  
XX (particularly chronic bronchitis) or diabetes (particularly type 1  
XX diabetes), and B-cell malignancies (claimed). They are also useful  
XX for studying interactions between CD23 and various ligands and  
XX determining the binding agents.  
XX  
XX Sequence 145 AA;  
SQ  
Query Match 100.0%; Score 31; DB 21; Length 145;  
Best Local Similarity 100.0%; Pred. NO. 8.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 86 LMSTRAS 92  
RESULT 13  
AAW39816  
ID AAW39816 standard; peptide; 7 AA.  
XX  
XX AAW39816;  
XX  
XX 16-JUN-1998 (first entry)  
XX  
XX Light chain CDR2 of catalytic antibody 3B9.  
XX  
XX Variable domain; lambda light chain; Catalytic antibody; degradation;  
XX cocaine; cocaine transition state analogue; TSA; benzoic acid;  
XX phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
XX overdose; addiction.  
XX  
XX Mus sp.  
XX  
XX WO9749800-A1.  
XX  
XX 31-DEC-1997.  
XX  
XX 25-JUN-1997; 97WO-US10965.  
XX  
XX 25-JUN-1996; 96US-0672345.  
XX  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX  
XX Landry DW;  
XX  
XX WPI; 1998-077166/07.  
XX  
XX New catalytic antibodies able to decompose cocaine, single-chain  
XX analogues - used to treat cocaine overdose and addiction, required  
XX in far smaller doses than antibodies that antagonise cocaine by  
XX simply binding  
XX Claim 11; Page 80; 147pp; English.  
XX  
XX AA39815-17 represent the sequences of the light chain complementarity  
XX determining regions (CDRs) of the catalytic antibody 3B9, which is able  
XX to degrade cocaine. A series of cocaine transition state analogues  
XX (TSAs) were prepared and used to immunise mice for production of  
XX hybridomas. Catalytic antibodies were identified by their capacity to  
XX release 3H-benzoic acid from 3H-phenyl cocaine. The 3B9 antibody was  
XX identified using TSA1, which is an immunogenic conjugate of a phosphate  
XX monoester transition state analogue. Antibody 3B9 has a per minute Kcat  
XX of 0.11. The antibodies reduce the concentration of cocaine in a subject,  
XX and are used particularly for the treatment of an overdose. They are also  
XX used for treating addiction (by reducing the in vivo concentration that  
XX can be achieved).  
XX  
XX Sequence 7 AA;  
SQ  
Query Match 90.3%; Score 28; DB 19; Length 7;  
Best Local Similarity 85.7%; Pred. NO. 7.8e+05;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMSTRAS 7  
DB 1 LMSTRSS 7  
RESULT 14  
AAW39801  
ID AAW39801 standard; protein; 113 AA.  
XX  
XX AAW39801;  
XX  
XX 16-JUN-1998 (first entry)  
XX  
XX Variable domain of the Kappa light chain of catalytic antibody 3B9.  
XX  
XX Variable domain; lambda light chain; catalytic antibody; degradation;  
XX cocaine; cocaine transition state analogue; TSA; benzoic acid;  
XX phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
XX overdose; addiction.  
XX  
XX Mus sp.  
XX  
XX WO9749800-A1.  
XX  
XX 31-DEC-1997.  
XX  
XX 25-JUN-1997; 97WO-US10965.  
XX  
XX 25-JUN-1996; 96US-0672345.  
XX  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX  
XX Landry DW;  
XX  
XX PI

XX WPI; 1998-077166/07.  
 DR P-PSDB; AAV09791.  
 XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX Claim 12; Pages 71-72; 147pp; English.  
 PS AAW39801-05 represent the amino acid sequences of the variable domain  
 CC of the kappa light chain of catalytic antibodies which are able to  
 CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
 CC were prepared and used to immunise mice for production of hybridomas.  
 CC Catalytic antibodies were identified by their capacity to release  
 CC 3H-benzoic acid from 3H-phenyl cocaine. The 389 antibody (AAW39806  
 CC represents the heavy chain) was identified using TSA1, which is an  
 CC immunogenic conjugate of a phosphate monoester transition state  
 CC analogue. Antibody 389 has a per minute Kcat of 0.11. The antibodies  
 CC reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used  
 CC for treating addiction (by reducing the in vivo concentration that can  
 CC be achieved).  
 XX SQ Sequence 113 AA;  
 Query Match 90.3%; Score 28; DB 19; Length 113;  
 Best Local Similarity 85.7%; Pred. No. 35;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LMSTRAS 7  
 Db 55 LMSTRSS 61  
 RESULT 15  
 AAW39899  
 ID AAW39899 standard; Protein; 274 AA.  
 XX AC AAW39899;  
 XX DT 16-JUN-1998 (first entry)  
 XX DE Single chain Fv region of the catalytic antibody 389.  
 KW Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX OS Mus sp.  
 FH Key Location/Qualifiers  
 FT Region 32..37 "complementarity determining region 1 of the  
 FT heavy chain"  
 FT Region 52..67 "complementarity determining region 2 of the  
 FT heavy chain"  
 FT Region 99..106 "complementarity determining region 3 of the  
 FT heavy chain"  
 FT Region 120..134 "complementarity determining region 1 of the  
 FT heavy chain"  
 FT Region 159..174 "complementarity determining region 2 of the  
 FT heavy chain"  
 FT Region 186..192 "complementarity determining region 3 of the  
 FT heavy chain"  
 FT Region 225..233 "complementarity determining region 1 of the  
 FT heavy chain"

FT Region 259..263 light chain"  
 FT /note= "Flag epitope sequence"  
 XX WO9749800-A1.  
 XX 31-DEC-1997.  
 XX 25-JUN-1997; 97WO-US10965.  
 XX 25-JUN-1996; 96US-0672345.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Landry DW;  
 XX WPI; 1998-077166/07.  
 XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX Disclosure; Fig 27; 147pp; English.  
 XX The present sequence represents the single chain Fv region of the  
 CC monoclonal catalytic antibody 389, which is capable of degrading  
 CC cocaine. A series of cocaine transition state analogues (TSAs) were  
 CC prepared and used to immunise mice for production of hybridomas.  
 CC Catalytic antibodies were identified by their capacity to release  
 CC 3H-benzoic acid from 3H-phenyl cocaine. The antibodies reduce the  
 CC concentration of cocaine in a subject, and are used particularly for  
 CC the treatment of an overdose. They are also used for treating addiction  
 CC (by reducing the in vivo concentration that can be achieved).  
 XX SQ Sequence 274 AA;  
 Query Match 90.3%; Score 28; DB 19; Length 274;  
 Best Local Similarity 85.7%; Pred. No. 91;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LMSTRAS 7  
 Db 186 LMSTRSS 192  
 Search completed: April 22, 2003, 12:51:12  
 Job time : 18.75 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:53:59 ; Search time 7.375 Seconds  
(without alignments)  
76.055 Million cell updates/sec

Title: US-09-674-716B-5  
Perfect score: 31  
Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pap.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pap.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pap.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pap.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pap.\*
- 6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pap.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pap.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pap.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pap.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pap.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pap.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US10\_PUBCOMB.pap.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pap.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	100	10	US-09-840-459-32 Sequence 32, Appl
2	26	83.9	115	10	US-09-864-761-33844 Sequence 33844, A
3	26	83.9	449	10	US-09-925-301-884 Sequence 884, App
4	26	83.9	810	9	US-09-712-363-281 Sequence 281, App
5	26	83.9	1464	9	US-10-060-036-159 Sequence 159, App
6	24	77.4	32	10	US-09-864-761-45487 Sequence 45487, A
7	24	77.4	108	10	US-09-764-877-1708 Sequence 1708, App
8	24	77.4	114	9	US-10-083-357-784 Sequence 784, App
9	24	77.4	274	9	US-09-866-050A-336 Sequence 336, App
10	24	77.4	295	9	US-09-829-936A-31 Sequence 31, Appl
11	24	77.4	295	9	US-10-005-057A-5 Sequence 5, Appl
12	24	77.4	305	9	US-09-738-626-4088 Sequence 4088, App
13	24	77.4	309	10	US-09-908-711-78 Sequence 78, Appl
14	24	77.4	433	10	US-09-925-302-691 Sequence 691, App
15	24	77.4	442	9	US-09-829-936A-16 Sequence 16, Appl
16	24	77.4	443	9	US-10-028-072-318 Sequence 318, App
17	24	77.4	443	9	US-10-121-049-318 Sequence 318, App
18	24	77.4	443	9	US-10-123-904-318 Sequence 318, App
19	24	77.4	443	9	US-10-123-904-318 Sequence 318, App

20	24	77.4	443	9	US-10-140-470-318 Sequence 318, App
21	24	77.4	443	9	US-10-175-746-318 Sequence 318, App
22	24	77.4	443	9	US-10-176-921-318 Sequence 318, App
23	24	77.4	443	9	US-10-176-921-318 Sequence 318, App
24	24	77.4	443	9	US-10-137-865-318 Sequence 318, App
25	24	77.4	443	9	US-10-140-474-318 Sequence 318, App
26	24	77.4	443	9	US-10-142-431-318 Sequence 318, App
27	24	77.4	443	9	US-10-143-114-318 Sequence 318, App
28	24	77.4	443	9	US-10-140-002-318 Sequence 318, App
29	24	77.4	443	9	US-10-142-419-318 Sequence 318, App
30	24	77.4	443	9	US-09-829-936A-22 Sequence 22, Appl
31	24	77.4	443	9	US-10-123-262-318 Sequence 318, App
32	24	77.4	443	9	US-10-142-423-318 Sequence 318, App
33	24	77.4	443	9	US-10-121-050-318 Sequence 318, App
34	24	77.4	443	9	US-10-141-755-318 Sequence 318, App
35	24	77.4	443	9	US-10-143-033-318 Sequence 318, App
36	24	77.4	443	9	US-10-123-108-318 Sequence 318, App
37	24	77.4	443	9	US-10-123-236-318 Sequence 318, App
38	24	77.4	443	9	US-10-123-261-318 Sequence 318, App
39	24	77.4	443	9	US-10-140-921-318 Sequence 318, App
40	24	77.4	443	9	US-10-140-928-318 Sequence 318, App
41	24	77.4	443	9	US-10-121-045-318 Sequence 318, App
42	24	77.4	443	9	US-10-123-292-318 Sequence 318, App
43	24	77.4	443	9	US-10-123-903-318 Sequence 318, App
44	24	77.4	443	9	US-10-124-819-318 Sequence 318, App
45	24	77.4	443	9	US-10-124-822-318 Sequence 318, App

## ALIGNMENTS

RESULT 1  
US-09-840-459-32  
; Sequence 32, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: Larosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Siobhan H.  
; APPLICANT: O'Keefe, Theresa  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; FILE REFERENCE: 1855.1052-012  
; CURRENT APPLICATION NUMBER: US/09/840,459  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 09/497,625  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 09/359,193  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 09/121,781  
; PRIOR FILING DATE: 1998-07-23  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 32  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-09-840-459-32

Query Match 100.0%; Score 31; DB 10; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7  
Db 55 LMSTRAS 61

RESULT 2



```
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-09-712-363-281

Query Match      83.9%; Score 26; DB 9; Length 810;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 38 LMSNRAS 44

RESULT 5
US-10-060-036-159
; Sequence 159, Application US/10060036
; Publication No. US200300731441
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-060-036-159

Query Match      83.9%; Score 26; DB 9; Length 1464;
Best Local Similarity 85.7%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 1357 LMSTRAS 1363

RESULT 6
US-09-864-761-45487
; Sequence 45487, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acamica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-877-1708
; Sequence 1708, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1708
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-877-1708

Query Match      77.4%; Score 24; DB 10; Length 108;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 1 LMSTRAS 7

RESULT 7
US-09-764-877-1708
; Sequence 1708, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1708
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-877-1708

Query Match      77.4%; Score 24; DB 10; Length 32;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MSTRAS 7
Db 2 LSTRAS 7

US-09-864-761-45487
; OTHER INFORMATION: MAP TO AC024196.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.45
; OTHER INFORMATION: EST_HUMAN HIT: AUI21127.1, EVALUE 1.50e-02
; US-09-864-761-45487

Query Match      77.4%; Score 24; DB 10; Length 32;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MSTRAS 7
Db 2 LSTRAS 7
```

```

Db      67 LPSTRAS 73

RESULT 8
US-10-083-357-784
; Sequence 784, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Qiangdong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083,357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 784
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-784

Query Match      77.4%; Score 24; DB 9; Length 114;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LMSTRAS 7
      |||||
Db     19 LMTTRGS 25

RESULT 9
US-09-866-050A-336
; Sequence 336, Application US/09866050A
; Publication No. US2003004071A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 10000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 336
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Human
US-09-866-050A-336

Query Match      77.4%; Score 24; DB 9; Length 274;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LMSTRAS 7
      |||||
Db     254 LMSYRAS 260

RESULT 10
US-09-829-936A-9
; Sequence 9, Application US/09829936A
; Publication No. US20030049699A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharma, S.A.
; TITLE OF INVENTION: Polypeptide (MBP1) Capable Of Interacting With Oncogenic Mutants
; FILE REFERENCE: ST98033
; CURRENT APPLICATION NUMBER: US/09/829,936A
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/251,555
; PRIOR FILING DATE: 1998-10-12
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 9
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Fragment C-term MBP1 murine
US-09-829-936A-9

Query Match      77.4%; Score 24; DB 9; Length 295;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LMSTRAS 7
      |||||
Db     275 LMSYRAS 281

RESULT 11
US-09-829-936A-31
; Sequence 31, Application US/09829936A
; Publication No. US20030049699A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharma, S.A.
; TITLE OF INVENTION: Polypeptide (MBP1) Capable Of Interacting With Oncogenic Mutants
; FILE REFERENCE: ST98033
; CURRENT APPLICATION NUMBER: US/09/829,936A
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR9812754
; PRIOR FILING DATE: 1998-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human fragment C-term MBP1
US-09-829-936A-31

Query Match      77.4%; Score 24; DB 9; Length 295;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LMSTRAS 7
      |||||
Db     275 LMSYRAS 281

RESULT 12
US-10-005-057A-5
; Sequence 5, Application US/10005057A
; Patent No. US20020170087A1
; GENERAL INFORMATION:
; APPLICANT: Tao Yumin
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Shen, Bo
; APPLICANT: Lowe, Keith S.
; APPLICANT: Danilevskaya, Olga
; APPLICANT: Mahajan, Pramod
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Sakai, Hajime
; APPLICANT: Klein, Ted M.
; TITLE OF INVENTION: Transcriptional Regulator Nucleic Acids,
; FILE REFERENCE: 1288
; TITLE OF INVENTION: Polypeptides and Methods of Use Thereof
; CURRENT APPLICATION NUMBER: US/10/005,057A
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/251,555
; PRIOR FILING DATE: 2001-12-04

```

Thu Apr 24 09:13:43 2003

us-09-674-716b-5.open.rapb

```
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(425)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-005-057A-5

Query Match          77.4%; Score 24; DB 9; Length 297;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRA 6
   |:|||
Db 66 LLSTRA 71

RESULT 13
US-09-738-626-4088
; Sequence 4088, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4088
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4088

Query Match          77.4%; Score 24; DB 9; Length 305;
Best Local Similarity 71.4%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
   |:|||
Db 246 VMSSRAS 252

RESULT 14
US-09-908-711-78
; Sequence 78, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA128
; CURRENT APPLICATION NUMBER: US/09/908,711
```

```

; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 78
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-908-711-78

```

```

Query Match      77.4%; Score 24; DB 10; Length 309;
Best Local Similarity 71.4%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LMSTRAS 7
|:|:|
Db 202 LLSRAS 208

```

# RESULT 15

```

US-09-925-302-691
; Sequence 691, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 691
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-691

```

```

Query Match      77.4%; Score 24; DB 10; Length 433;
Best Local Similarity 85.7%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LMSTRAS 7
|:|:|
Db 413 LMSYRAS 419

```

Search completed: April 22, 2003, 13:11:12  
Job time : 8.375 secs





J. Immunol. 141, 4012-4019, 1988  
 A:Title: Genetics of the phosphocholine-specific antibody response to Streptococcus pneumoniae  
 A:Reference number: A30534; UID:89035545; PMID:3141511  
 A:Accession: G30538  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-74 <CLA>  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 87.5%; Score 42; DB 2; Length 74;  
 Best Local Similarity 88.9%; Pred. No. 0.16;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
 |||||  
 Db 63 QQLVEYPLT 71

RESULT 3  
 KMS16  
 Ig kappa chain V region (M167) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 30-Nov-1980 #sequence\_revision 30-Nov-1980 #text\_change 16-Aug-1996  
 C:Accession: A01908  
 R:Rudikoff, S.; Potter, M.  
 Biochemistry 17, 2703-2707, 1978  
 A:Title: Kappa-Chain variable region from M167, a phosphorylcholine binding myeloma protein  
 A:Reference number: A01908; UID:7900273; PMID:99160  
 A:Accession: A01908  
 A:Molecule type: protein  
 A:Residues: 1-112 <RUD>  
 C:Comment: This chain was isolated from a myeloma protein that binds phosphorylcholine.  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into larger disulfide bonds.  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:16-95/Domain: immunoglobulin homology <IMM>  
 F:23-93/Disulfide bonds: #status predicted

Query Match 87.5%; Score 42; DB 1; Length 112;  
 Best Local Similarity 88.9%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
 |||||  
 Db 94 QQLVEYPLT 102

RESULT 4  
 KMS51  
 Ig kappa chain V region (M511) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 21-Jan-2000  
 C:Accession: A01910  
 R:Appella, E.  
 Mol. Immunol. 17, 711-718, 1980  
 A:Title: Amino acid sequence of the light chain variable region of M511, a phosphorylcholine binding myeloma protein  
 A:Reference number: A01910; UID:81052016; PMID:6776396  
 A:Accession: A01910  
 A:Molecule type: protein  
 A:Residues: 1-113 <APP>  
 C:Comment: This chain was isolated from a myeloma protein that binds phosphorylcholine.  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into larger disulfide bonds.  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:16-95/Domain: immunoglobulin homology <IMM>  
 F:23-93/Disulfide bonds: #status predicted

Query Match 87.5%; Score 42; DB 1; Length 113;  
 Best Local Similarity 88.9%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
 |||||  
 Db 94 QQLVEYPLT 102

RESULT 5  
 KMS67  
 Ig kappa chain precursor V region (VK167) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 21-Jan-2000  
 C:Accession: A01909  
 R:Selsing, E.; Storb, U.  
 Cell 25, 47-58, 1981  
 A:Title: Somatic mutation of immunoglobulin light-chain variable-region genes.  
 A:Reference number: A01909; UID:8200223; PMID:6791832  
 A:Accession: A01909  
 A:Molecule type: DNA  
 A:Residues: 1-120 <SEL>  
 A:Note: the sequence was determined from the germline gene  
 C:Genetics:  
 A:Introns: 17/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into larger disulfide bonds.  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-120/Product: Ig kappa chain V region (VK167) #status predicted <MAT>  
 F:36-113/Domain: immunoglobulin homology <IMM>  
 F:43-113/Disulfide bonds: #status predicted

Query Match 77.1%; Score 37; DB 1; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 2.8;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQLVEYP 7  
 |||||  
 Db 114 QQLVEYP 120

RESULT 6  
 A85363  
 probable calmodulin-binding protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: A85363  
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory  
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A:Reference number: A85001; UID:20083488; PMID:10617198  
 A:Accession: A85363  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-467 <SPO>  
 A:Cross-references: GB:NC\_001268; NID:g7270002; PIDN:CAB79818.1; GSPDB:GN00140  
 C:Genetics:  
 A:Gene: AT4g31000  
 A:Map position: 4

Query Match 77.1%; Score 37; DB 2; Length 467;  
 Best Local Similarity 75.0%; Pred. No. 12;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQLVEYPF 8  
 |||||  
 Db 334 QHLIEYFP 341

RESULT 7  
 S36277  
 Ig lambda chain V region (clone alpha-FOG1-G8) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 21-Jan-2000

C;Accession: S36277  
 R;Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.  
 EMBO J. 12, 725-734, 1993  
 A;Title: Human anti-self antibodies with high specificity from phage display libraries.  
 A;Reference number: S36256; MUID:93178448; PMID:7679990  
 A;Accession: S36277  
 A;Status: preliminary; nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-108 <GRI>  
 A;Cross-references: EMBL:Z18829; NID:g33417; PIDN:CAA79281.1; PID:g939910  
 C;Superfamily: immunoglobulin V region; immunoglobulin homology  
 C;Keywords: heterotetramer; immunoglobulin  
 F;16-90/Domain: immunoglobulin homology <IMM>

Query Match 75.0%; Score 36; DB 2; Length 108;  
 Best Local Similarity 66.7%; Pred. No. 4;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPT 9  
 |||: |||  
 Db 89 QQLISYPLT 97

RESULT 8  
 T38148  
 phosphatidyl synthase - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C;Accession: T38148  
 R;Pearson, D.; Church, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, September 1997  
 A;Reference number: Z21774  
 A;Accession: T38148  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-570 <PEA>  
 A;Cross-references: EMBL:Z99295; PIDN:CAB16578.1; GSPDB:GN00066; SPDB:SPAC22A12.08C  
 A;Experimental source: strain 972h-; cosmid c22A12  
 C;Genetics:  
 A;Gene: SPDB:SPAC22A12.08C  
 A;Map position: 1  
 A;Introns: 43/2; 62/2; 227/2; 483/1

Query Match 75.0%; Score 36; DB 2; Length 570;  
 Best Local Similarity 77.8%; Pred. No. 23;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPT 9  
 |||: |||  
 Db 547 QQLVEYSP 555

RESULT 9  
 B29775  
 Ig kappa chain precursor V region (mouse 24.1) - shrew mouse  
 C;Species: Mus pahari  
 C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 21-Jan-2000  
 C;Accession: B29775  
 R;Jouvin-Marche, E.; Rudikoff, S.  
 Immunogenetics 24, 191-201, 1986  
 A;Title: Evolution of a V-kappa gene family.  
 A;Reference number: A91751; MUID:87006895; PMID:3093373  
 A;Accession: B29775  
 A;Molecule type: DNA  
 A;Residues: 1-120 <JOU>  
 A;Cross-references: GB:ML553; NID:g197470; PIDN:AAA39037.1; PID:g197471  
 A;Note: this sequence was determined from the germline gene  
 C;Genetics:  
 A;Introns: 17/1  
 C;Superfamily: immunoglobulin V region; immunoglobulin homology  
 C;Keywords: heterotetramer; immunoglobulin  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;21-120/Product: Ig kappa chain V region 24.1 #status predicted <MAT>

F;36-115/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 72.9%; Score 35; DB 2; Length 120;  
 Best Local Similarity 85.7%; Pred. No. 7.2;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQLVEYP 7  
 |||: |||  
 Db 114 QQLVEYP 120

## RESULT 10

A86371

hypothetical protein F508.25 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C;Accession: A86371

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A86371

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-480 &lt;STO&gt;

A;Cross-references: GB:AE005172; NID:g4056452; PIDN:AAC98025.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

## Query Match

Best Local Similarity 72.9%; Score 35; DB 2; Length 480;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QQLVEYFP 8

|||: |||

Db 269 QQLVEYFP 275

## RESULT 11

D82100

mesJ protein VC2242 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C;Accession: D82100

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.

L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82100

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-440 &lt;HEI&gt;

A;Cross-references: GB:AE004296; GB:AE003852; NID:g9656799; PIDN:AAF95386.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC2242

A;Map position: 1

C;Superfamily: hypothetical protein HI0404

## Query Match

Best Local Similarity 70.8%; Score 34; DB 2; Length 440;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYFP 8

Db 10 QQLARYPF 17  
||| |||

## RESULT 12

D89903  
hypothetical protein SAIL131 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: D89903  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino, A.; Mizukami, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: D89903  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-586 <KUR>  
A:Cross-references: GB:BA000018; PID:gl3701089; PIDN:BAB42384.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Gene: SAIL131  
C:Superfamily: Halobacterium halobium 2-oxoacid ferredoxin oxidoreductase; 2-oxoacid ferredoxin

Query Match 70.8%; Score 34; DB 2; Length 586;  
Best Local Similarity 75.0%; Pred. No. 62;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QLVVEYPT 9  
||| |||

Db 109 QQLARYPT 116

## RESULT 13

S76367  
hypothetical protein - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis sp.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S76367  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-722 <KAN>  
A:Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAA10219.1; PID:g100159  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: glycine-tRNA ligase beta chain

Query Match 70.8%; Score 34; DB 2; Length 722;  
Best Local Similarity 85.7%; Pred. No. 78;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQLVEYPT 7  
||| |||

Db 254 QQLVEYPT 260

## RESULT 14

T01906  
hypothetical protein T12H20.3 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 24-Mar-1999  
C:Accession: T01906  
R:Corton, M.; Graves, T.; Sutterer, C.; Modde, T.  
submitted to the EMBL Data Library, July 1998

A:Description: The sequence of A. thaliana T12H20.  
A:Reference number: Z14453  
A:Accession: T01906  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1074 <COT>  
A:Cross-references: EMBL:AF080119; NID:g3600029; PID:g3600033  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Map position: 4  
A:Note: T12H20.3

Query Match 70.8%; Score 34; DB 2; Length 1074;  
Best Local Similarity 71.4%; Pred. No. 12e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QLVVEYPT 8  
||| |||

Db 853 QLVVEYPT 859

## RESULT 15

I26317  
IG kappa chain V region (H158-89H4) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 21-Jan-2000  
C:Accession: I26317  
R:Caton, A.J.; Brownlee, G.G.; Staudt, L.M.; Gerhard, W.  
EMBO J. 5, 1577-1587, 1986  
A:Title: Structural and functional implications of a restricted antibody response to a d  
A:Reference number: A91043; MUID:86300658; PMID:2427335  
A:Accession: I26317  
A:Molecule type: DNA  
A:Residues: 1-112 <CAT>  
A:Experimental source: strain Balb/c  
A:Note: this sequence was determined from the germline gene  
C:Comment: This chain was isolated from a hybridoma protein that binds influenza virus h  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:16-95/Domain: immunoglobulin homology <IMM>

Query Match 68.8%; Score 33; DB 2; Length 112;  
Best Local Similarity 75.0%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QLVVEYPT 9  
||| |||

Db 95 QLVVEYPT 102

Search completed: April 22, 2003, 12:54:49  
Job time : 10.8393 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

QM protein - protein search, using sw model

Run on: April 22, 2003, 12:46:53 ; Search time 4.5 Seconds  
(without alignments)  
82.953 Million cell updates/sec

Title: US-09-674-716B-7

Perfect score: 48

Sequence: 1 QQLVEVPFT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	89.6	383	1	CYB APILI
2	42	87.5	112	1	KV2A MOUSE
3	42	87.5	113	1	KV2C MOUSE
4	37	77.1	120	1	KV2B MOUSE
5	36	75.0	379	1	CYB AKOJE
6	36	75.0	379	1	CYB AKOTO
7	36	75.0	381	1	CYB SIGHI
8	34	70.8	722	1	SYGB SYNXY3
9	33	68.8	452	1	CN17_DICDI
10	33	68.8	561	1	LCFA_ECOLI
11	33	68.8	664	1	SYGB_AQUAE
12	33	68.8	792	1	ATX1_MOUSE
13	33	68.8	816	1	ATX1_HUMAN
14	32	66.7	295	1	BIEA RAT
15	32	66.7	296	1	BIEA HUMAN
16	32	66.7	374	1	Y006 BORBU
17	32	66.7	379	1	CYB BOLAM
18	32	66.7	673	1	2145 HUMAN
19	31	64.6	130	1	MSRR_PASMU
20	31	64.6	168	1	CRACA
21	31	64.6	178	1	YP20_BACLI
22	31	64.6	222	1	TYSY_METJA
23	31	64.6	366	1	HNSL_SORBI
24	31	64.6	379	1	CYB ALCAA
25	31	64.6	379	1	CYB AMBCI
26	31	64.6	379	1	CYB AMWHA
27	31	64.6	379	1	CYB ARGZ
28	31	64.6	379	1	CYB CAPCA
29	31	64.6	379	1	CYB CAPPY
30	31	64.6	379	1	CYB CEREL
31	31	64.6	379	1	CYB CERNI
32	31	64.6	379	1	CYB CHEMA
33	31	64.6	379	1	CYB DAMDA

34 31 64.6 379 1 CYB\_EUTDO  
35 31 64.6 379 1 CYB\_GLASA  
36 31 64.6 379 1 CYB\_HYDIN  
37 31 64.6 379 1 CYB\_SAISC  
38 31 64.6 379 1 CYB\_SCINI  
39 31 64.6 379 1 CYB\_SPEAR  
40 31 64.6 379 1 CYB\_SPEMO  
41 31 64.6 379 1 CYB\_SPERI  
42 31 64.6 380 1 CYB\_HUMAN  
43 31 64.6 380 1 CYB\_LATCH  
44 31 64.6 380 1 CYB\_MICAR  
45 31 64.6 380 1 CYB\_MICLO

Q9tf32 eutamias do  
O48372 glaucomyx s  
O47930 hydropotes  
Q35930 saimiri sci  
Q35895 sciurus nig  
Q9tf51 spermophilu  
Q9tf31 spermophilu  
P49341 spermophilu  
P00156 homo sapien  
O03176 latimeria c  
Q36922 microtus ar  
P56731 microtus lo

## ALIGNMENTS

RESULT 1  
CYB APILI STANDARD; PRT; 383 AA.  
AC P34845;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Cytochrome B.  
GN COB OR CYTB.  
OS Apis mellifera ligustica (Common honeybee).  
OG Mitochondrion.  
OC Insecta; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;  
OC Aculeata; Apoidea; Apidae; Apis.  
OX NCBI\_TaxID:7469;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thorax;  
RX MEDLINE=92261310; PubMed=1533894;  
RA Crozier R.H., Crozier Y.C.;  
RT "The cytochrome b and ATPase genes of honeybee mitochondrial DNA.";  
Mol. Biol. Evol. 9:474-482(1992).  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thorax;  
RX MEDLINE=93114603; PubMed=8417993;  
RA Crozier R.H., Crozier Y.C.;  
RT "The mitochondrial genome of the honeybee Apis mellifera: complete  
sequence and genome organization.";  
Genetics 133:97-117(1993).  
CC -I- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE  
COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A  
RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL  
COUPLED TO ATP SYNTHESIS.  
CC -I- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
BOUND TO THE PROTEIN.  
CC -I- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
CYTOCHROME C1 AND THE RIESKE PROTEIN.  
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L06178; AAB96809.1; -  
CC EMBL: M87052; -; NOT\_ANNOTATED\_CDS.  
CC PIR: C43622; C43622.  
CC InterPro: IPR000179; Cyt\_b\_b6.  
CC Pfam: PF00032; cytochrome\_b\_c1.  
CC Pfam: PF00032; cytochrome\_b\_c1.  
CC PROSITE: PS00192; CYTOCHROME\_B\_HEME; 1.  
CC PROSITE: PS00193; CYTOCHROME\_B\_QQ; 1.

```
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 85 85 IRON 1 (HEME B562 AXIAL LIGAND) .
FT METAL 99 99 IRON 2 (HEME B566 AXIAL LIGAND) .
FT METAL 184 184 IRON 1 (HEME B562 AXIAL LIGAND) .
FT METAL 198 198 IRON 1 (HEME B566 AXIAL LIGAND) .
SQ SEQUENCE 383 AA; 45256 MW; A140A05B6053C2D5 CRC64;

Query Match 89.6%; Score 43; DB 1; Length 383;
Best Local Similarity 77.8%; Pred. No. 0.32;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9
DB 341 QQLIEYPPT 349

RESULT 2
KV2A_MOUSE STANDARD; PRT; 112 AA.
AC P01626;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region MOPC 167.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SQ SEQUENCE 112 AA; 12349 MW; A58EDF56404B9726 CRC64;

Query Match 87.5%; Score 42; DB 1; Length 112;
Best Local Similarity 88.9%; Pred. No. 0.15;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9
DB 94 QQLVEYPLT 102

RESULT 3
KV2C_MOUSE STANDARD; PRT; 113 AA.
AC P01628;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
```

```
DE Ig kappa chain V-II region MOPC 511.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SQ SEQUENCE.
RX MEDLINE=81052016; PubMed=6776396;
RA Appella E.;
RT "Amino acid sequence of the light chain variable region of M511, a
phosphorylcholine-binding murine myeloma protein.";
RL Mol. Immunol. 17:711-718 (1980).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
BINDS PHOSPHORYLCHOLINE.
DR.
DR PIR; A01910; KVM51.
DR HSP; P80362; IWT.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 39
FT DOMAIN 40 54
FT DOMAIN 55 61
FT DOMAIN 62 93
FT DOMAIN 94 102
FT DOMAIN 103 112
FT DISULFID 23 93
FT NON_TER 112 112
SQ SEQUENCE 113 AA; 12496 MW; EFB0DC4DA2BD3450 CRC64;

Query Match 87.5%; Score 42; DB 1; Length 113;
Best Local Similarity 88.9%; Pred. No. 0.15;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9
DB 94 QQLVEYPLT 102

RESULT 4
KV2B_MOUSE STANDARD; PRT; 120 AA.
AC P01627;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region VKAPPA167 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SQ SEQUENCE FROM N.A.
RX MEDLINE=82002223; PubMed=6791832;
RA Selsing E.; Storb U.;
RT "Somatic mutation of immunoglobulin light-chain variable-region
genes.";
RL Cell 25:47-58 (1981).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC
CC EMBL; J00562; AAA39032.1; -
CC EMBL; K02415; AAA39051.1; -
DR PIR; A01909; KVM567.
DR HSP; P80362; IWT.
```

DR InterPro; IPR003006; Ig MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
KW Immunoglobulin V region; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 120 IG KAPPA CHAIN V-II REGION VKAPPAL67.  
FT DOMAIN 21 43 FRAMEWORK-1.  
FT DOMAIN 21 43 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 44 59 FRAMEWORK-2.  
FT DOMAIN 60 74 FRAMEWORK-3.  
FT DOMAIN 75 81 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 82 113 FRAMEWORK-3.  
FT DOMAIN 82 113 COMPLEMENTARITY-DETERMINING-3.  
FT DISULFID 114 120 BY SIMILARITY.  
FT SEQUENCE 120 AA; 13280 MW; 63BB571F0E4DE3E8 CRC64;  
SQ

Query Match 77.1%; Score 37; DB 1; Length 120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QQLVEYP 7  
Db 114 QQLVEYP 120

RESULT 5  
CYB AKOJE  
ID CYB AKOJE STANDARD; PRT; 379 AA.  
AC P21715;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Cytochrome B.  
GN MTCYB OR COB OR CYTB.  
OS Akodon jelskii.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
OC Akodon.  
OX NCBI\_TaxID=10079;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Smith M.F.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE OF 1-267 FROM N.A.  
RC TISSUE=Liver;  
RA Smith M.F., Patton J.L.;  
RT "The diversification of South American murid rodents: evidence from mitochondrial DNA sequence data for the akodontine tribe.";  
RL Biol. J. Linn. Soc. Lond. 50:149-177(1993).  
[3]  
RN SEQUENCE OF 1-133 FROM N.A.  
RP STRAIN=mvz catalog 173073, 173074, 173083, and 173084; TISSUE=Liver;  
RC MEDLINE=91163325; PubMed=2002767;  
RA Smith M.F., Patton J.L.;  
RT "Variation in mitochondrial cytochrome b sequence in natural populations of South American akodontine rodents (Muridae: Sigmodontinae).";  
RL Mol. Biol. Evol. 8:85-103(1991).  
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS.  
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN.  
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC  
CC  
DR EMBL; M35714; AAA16998.2; ALT\_TERM.  
DR EMBL; M35715; AAA31630.1; -.  
DR EMBL; M35716; AAA31629.1; -.  
DR PIR; C41824; C41824.  
DR InterPro; IPR000179; Cyt\_b\_b6.  
DR Pfam; PF00032; cytochrome b\_c; 1.  
DR Pfam; PF00033; cytochrome b\_n; 1.  
DR PROSITE; PS00193; CYTOCHROME B\_QQ; 1.  
DR PROSITE; PS00192; CYTOCHROME B\_HEME; 1.  
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;  
KW Heme.  
FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND).  
FT METAL 97 97 IRON 2 (HEME B566 AXIAL LIGAND).  
FT METAL 182 182 IRON 2 (HEME B562 AXIAL LIGAND).  
FT METAL 196 196 IRON 1 (HEME B566 AXIAL LIGAND).  
FT VARIANT 118 118 V -> I (IN STRAINS 173083 AND 173084).  
FT VARIANT 122 122 T -> A (IN STRAINS 173083 AND 173084).  
SQ SEQUENCE 379 AA; 42529 MW; F9F012A46671D59A CRC64;  
Query Match 75.0%; Score 36; DB 1; Length 379;  
Best Local Similarity 87.5%; Pred. No. 8;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 QQLVEYPFT 9  
Db 341 QPVEYPFT 348

RESULT 6  
CYB AKOTO  
ID CYB AKOTO STANDARD; PRT; 379 AA.  
AC P21721;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Cytochrome B.  
GN MTCYB OR COB OR CYTB.  
OS Akodon torques.  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
OC Akodon.  
OX NCBI\_TaxID=10078;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Smith M.F.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE OF 1-267 FROM N.A.  
RC TISSUE=Liver;  
RA Smith M.F., Patton J.L.;  
RT "The diversification of South American murid rodents: evidence from mitochondrial DNA sequence data for the akodontine tribe.";  
RL Biol. J. Linn. Soc. Lond. 50:149-177(1993).  
[3]  
RN SEQUENCE OF 1-133 FROM N.A.  
RP STRAIN=mvz catalog 171720, 171721, 174053, and 174054; TISSUE=Liver;  
RC MEDLINE=91163325; PubMed=2002767;  
RA Smith M.F., Patton J.L.;  
RT "Variation in mitochondrial cytochrome b sequence in natural populations of South American akodontine rodents (Muridae: Sigmodontinae).";  
RL Mol. Biol. Evol. 8:85-103(1991).  
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

CC COUPLED TO ATP SYNTHESIS.  
 CC COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
 CC BOUND TO THE PROTEIN.  
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M35700; AAA1897.2; -;  
 CC EMBL; M35701; AAA31624.1; -;  
 CC EMBL; M35702; AAA31625.1; -;  
 CC PIR; G23725; G23725.  
 CC InterPro; IPR000179; Cyt b b6.  
 CC Pfam; PF00032; cytochrome\_b\_c1.  
 CC Pfam; PF00033; cytochrome\_b\_n; 1.  
 CC PROSITE; PS00193; CYTOCHROME\_B\_QQ; 1.  
 CC PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
 CC Electron transport; Mitochondrion; Respiratory chain; Transmembrane;  
 CC Heme.  
 CC METAL 83 IRON 1 (HEME B562 AXIAL LIGAND).  
 CC METAL 97 IRON 2 (HEME B566 AXIAL LIGAND).  
 CC METAL 182 IRON 2 (HEME B562 AXIAL LIGAND).  
 CC METAL 196 IRON 1 (HEME B566 AXIAL LIGAND).  
 CC VARIANT 108 108 V -> A (IN STRAINS 174053 AND 174054).  
 CC VARIANT 115 115 V -> I (IN STRAINS 174053 AND 174054).  
 CC SEQUENCE 379 AA; 42611 MW; 4C94CF8A92CCC34F CRC64;  
 CC  
 CC Query Match 75.0%; Score 36; DB 1; Length 379;  
 CC Best Local Similarity 87.5%; Pred. No. 8;  
 CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 2 QLVYEPFT 9  
 CC | |||||  
 CC Db 341 QPVEYPT 348  
 CC  
 CC RESULT 7  
 CC CYB\_SIGHI STANDARD; PRT; 381 AA.  
 CC AC Q9XNU6;  
 CC DT 15-JUN-2002 (Rel. 41, Created)  
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Cytochrome B.  
 CC GN MTCYB OR COB OR CYTB.  
 CC OS Sigmmodon hispidus (Hispid cotton rat).  
 CC OG Mitochondrion.  
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmmodontinae;  
 CC OC Sigmmodon.  
 CC NCBI\_TaxID=42415;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC Smith M.P., Patton J.L.;  
 CC RT "Phylogenetic relationships and the radiation of sigmodontine rodents  
 CC in South America: evidence from cytochrome b.";  
 CC J. Mammal. Evol. 6:89-128(1999).  
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE  
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A  
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL  
 CC COUPLED TO ATP SYNTHESIS.  
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
 CC BOUND TO THE PROTEIN.  
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF108702; AAD45484.1; -;  
 CC InterPro; IPR000179; Cyt b b6.  
 CC Pfam; PF00032; cytochrome\_b\_c1.  
 CC Pfam; PF00033; cytochrome\_b\_n; 1.  
 CC PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
 CC PROSITE; PS00193; CYTOCHROME\_B\_QQ; 1.  
 CC Electron transport; Mitochondrion; Respiratory chain; Transmembrane;  
 CC Heme.  
 CC METAL 83 IRON 1 (HEME B562 AXIAL LIGAND).  
 CC METAL 97 IRON 2 (HEME B566 AXIAL LIGAND).  
 CC METAL 182 IRON 2 (HEME B562 AXIAL LIGAND).  
 CC METAL 196 IRON 1 (HEME B566 AXIAL LIGAND).  
 CC SEQUENCE 381 AA; 42965 MW; 0472259421B38284 CRC64;  
 CC  
 CC Query Match 75.0%; Score 36; DB 1; Length 381;  
 CC Best Local Similarity 87.5%; Pred. No. 8;  
 CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 2 QLVYEPFT 9  
 CC | |||||  
 CC Db 341 QPVEYPT 348  
 CC  
 CC RESULT 8  
 CC SYGB\_SYNY3 STANDARD; PRT; 722 AA.  
 CC AC Q55690;  
 CC DT 01-NOV-1997 (Rel. 35, Created)  
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Glycyl-tRNA synthetase beta chain (EC 6.1.1.14) (Glycine--tRNA ligase  
 CC beta chain) (GLYRS).  
 CC GN GLYS OR SLR0220.  
 CC OS Synechocystis sp. (strain PCC 6803).  
 CC OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 CC NCBI\_TaxID=1148;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE=96127529; PubMed=8590279;  
 CC Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
 CC Sugita M., Tabata S.;  
 CC "Sequence analysis of the genome of the unicellular cyanobacterium  
 CC Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
 CC region from map positions 64k to 92k of the genome.";  
 CC DNA Res. 2:153-166(1995).  
 CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) = AMP + diphosphate  
 CC + Glycyl-tRNA(Gly).  
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D64000; BAA10219.1; -;  
 CC InterPro; IPR002311; tRNA\_synth\_2f.  
 CC Pfam; PF02092; tRNA\_synth\_2f; 1.  
 CC PRINTS; PR01045; TRNASYNTHGB.



DR TIGRFAMs; TIGR00211; glyS; 1.  
 DR PROSITE; PS00861; AA TRNA LIGASE II GLYAB; 1.  
 KW Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding;  
 SQ SEQUENCE 722 AA; 80051 MW; F7085BA5A23436D7 CRC64;  
 Query Match 70.8%; Score 34; DB 1; Length 722;  
 Best Local Similarity 85.7%; Pred. No. 38;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QQLVEYP 7  
 :|||||  
 Db 254 EQLVEYP 260  
 :|||||  
 RESULT 9  
 ID\_CN17 D1CDI STANDARD; PRT; 452 AA.  
 AC P12019;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 3',5'-cyclic-nucleotide phosphodiesterase precursor (EC 3.1.4.17)  
 DE (PDEase) (3':5'-CNP).  
 GN PDSA.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87057386; PubMed=303365;  
 RA Lacombe M.-L., Podgorzki G.J., Franke J., Kessin R.H.;  
 RT "Molecular cloning and developmental expression of the cyclic  
 RT nucleotide phosphodiesterase gene of Dictyostelium discoideum.";  
 RL J. Biol. Chem. 261:16811-16817(1986).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89384622; PubMed=2779573;  
 RA Podgorzki G.J., Franke J., Faure M., Kessin R.H.;  
 RT "The cyclic nucleotide phosphodiesterase gene of Dictyostelium  
 RT discoideum utilizes alternate promoters and splicing for the  
 RT synthesis of multiple mRNAs.";  
 RL Mol. Cell. Biol. 9:3938-3950(1989).  
 [3]  
 RP SEQUENCE OF 1-69 FROM N.A.  
 RX MEDLINE=87010528; PubMed=3020155;  
 RA Podgorzki G.J., Franke J., Kessin R.H.;  
 RT "Isolation of a cDNA encoding a portion of the cyclic nucleotide  
 RT phosphodiesterase of Dictyostelium discoideum.";  
 RL J. Gen. Microbiol. 132:1043-1050(1986).  
 CC -1- FUNCTION: TO MAINTAIN THE RESPONSIVENESS OF CELLS TO THE  
 CC CHEMOTACTIC CAMP DURING THE AGGREGATION PHASE OF DEVELOPMENT.  
 CC -1- CATALYTIC ACTIVITY: Nucleoside 3',5'-cyclic phosphate + H(2)O =  
 CC nucleoside 5'-phosphate.  
 CC -1- SIMILARITY: BELONGS TO THE CAMP PHOSPHODIESTERASE CLASS-II FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; J02628; AAA68447.1; -.  
 DR EMBL; M23449; AAA63168.1; -.  
 DR EMBL; M15738; AAA33238.1; -.  
 DR PIR; A25346; A25346.  
 DR PIR; A32573; A32573.  
 DR DictyDb; DD02009; pdsa.  
 DR InterPro; IPR000396; Pdiesterase2.  
 DR Pfam; PF02112; PDEase II; 1.  
 DR PRINTS; PR00388; PDIESTERASE2.

DR ProDom; PD010003; Pdiesterase2; 2.  
 DR PROSITE; PS00607; PDEASE II; 1.  
 KW Glycoprotein; Hydrolase; CAMP; Zinc; Signal.  
 FT SIGNAL 1 23  
 FT PROPEP 24 49  
 FT CHAIN 50 452  
 FT 3',5'-CYCLIC-NUCLEOTIDE  
 FT PHOSPHODIESTERASE.  
 FT CARBOHYD 101 101  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 277 277  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 68 69  
 FT NW -> LT (IN REF. 3).  
 SQ SEQUENCE 452 AA; 51093 MW; A8F3C190D4603BD1 CRC64;  
 Query Match 68.8%; Score 33; DB 1; Length 452;  
 Best Local Similarity 83.3%; Pred. No. 38;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 VEYPFT 9  
 :|||||  
 Db 266 IEYPFT 271  
 :|||||  
 RESULT 10  
 ID\_CN17 D1CDI STANDARD; PRT; 561 AA.  
 AC P29212;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3) (Long-chain acyl-CoA  
 DE synthetase).  
 GN FADD OR OLDD OR B1805.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 EX MEDLINE=94150456; PubMed=8107670;  
 RA Fulda M., Heinz E., Wolter F.P.;  
 RT "The fadD gene of Escherichia coli K12 is located close to rnd at  
 RT 39.6 min of the chromosomal map and is a new member of the  
 RT AMP-binding protein family.";  
 RL Mol. Gen. Genet. 242:241-249(1994).  
 [2]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.  
 RC STRAIN=K12;  
 RX MEDLINE=93094273; PubMed=1460045;  
 RA Black P.N., Dirusso C.C., Metzger A.K., Heimert T.L.;  
 RT "Cloning, sequencing, and expression of the fadD gene of Escherichia  
 RT coli encoding acyl coenzyme A synthetase.";  
 RL J. Biol. Chem. 267:25513-25520(1992).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97251358; PubMed=9097040;  
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,  
 RA Isono K., Kasai H., Mizuro S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,  
 RA Saito N., Sampei G., Seki Y., Sivasubram S., Tagami H.,  
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;

RT "A 460-kb DNA sequence of the *Escherichia coli* K-12 genome  
 RL corresponding to the 40.1-50.0 min region on the linkage map."  
 CC DNA Res. 3:379-392 (1996).  
 CC -1- FUNCTION: ESTERIFICATION, CONCOMITANT WITH TRANSPORT, OF EXOGENOUS  
 CC LONG-CHAIN FATTY ACIDS INTO METABOLICALLY ACTIVE COA THIOESTERS  
 CC FOR SUBSEQUENT DEGRADATION OR INCORPORATION INTO PHOSPHOLIPIDS.  
 CC -1- CATALYTIC ACTIVITY: ATP + a long-chain carboxylic acid + CoA = AMP  
 CC + diphosphate + an acyl-CoA.  
 CC -1- COPACTOR: MAGNESIUM.  
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -1- SUBCELLULAR LOCATION: PARTIALLY MEMBRANE-ASSOCIATED (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X70994; CA450321.1; -;  
 DR EMBL; L02649; AAA23752.1; -;  
 DR EMBL; A5000275; AAC74875.1; -;  
 DR EMBL; D90823; BAA15600.1; -;  
 DR EMBL; D90824; BAA15609.1; -;  
 DR EMBL; D90825; BAA15614.1; -;  
 DR PIR; A45062; A45062.  
 DR HSP; P08659; 1LCI.  
 DR EcoGene; EG11530; fadd.  
 DR InterPro; IPR000873; AMP-bind.  
 DR Pfam; PF0501; AMP-binding; 1.  
 DR PROSITE; PS00455; AMP BINDING; 1.  
 DR KW Ligase; Fatty acid metabolism; Magnesium; Membrane; Complete proteome.  
 FT CONFLICT 34 51 ARYADQAPAFVNVNVEVMTF -> GALRRSTVCVEYGGNDL  
 (IN REF. 2).  
 FT CONFLICT 468 490 NEIEDVVMQHPQVEVRAVGVP ->  
 TELKSSCSMWAYKSRLLAYLP (IN REF. 2).  
 FT CONFLICT 496 496 A -> G (IN REF. 2).  
 FT CONFLICT 555 561 GKVDNKA -> QSGQ (IN REF. 2).  
 SQ SEQUENCE 561 AA; 62332 MW; 249B0AA54B3DBFA5 CRC64;  
 Query Match 68.8%; Score 33; DB 1; Length 561;  
 Best Local Similarity 55.6%; Pred. No. 47;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QQLVEYVPT 9  
 : : : : :  
 DB 293 KELAKVPT 301  
 RESULT 11  
 ID SYGB\_AQUAE STANDARD; PRT; 664 AA.  
 AC O67896;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Glycyl-tRNA synthetase beta chain (EC 6.1.1.14) (Glycine--tRNA ligase  
 DE beta chain) (GLYRS).  
 GN GLYS OR AQ.2141.  
 OS Aquifex aolicus.  
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;  
 CC Aquifex.  
 CC NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Qackerk G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RL aolicus";  
 CC Nature 392:353-358 (1998).  
 CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA (Gly) = AMP + diphosphate  
 CC + glycyl-tRNA (Gly).  
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains  
 CC (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000775; AAC07870.1; -;  
 DR InterPro; IPR002311; tRNA synt 2f.  
 DR Pfam; PF02092; tRNA synt 2f; 1.  
 DR PRINTS; PR01045; TRNASYNTHGB.  
 DR TIGRFAMs; TIGR00211; GLYS, 1.  
 DR PROSITE; PS50861; AA-tRNA\_LIGASE II GLYAB; 1.  
 DR KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 SQ SEQUENCE 664 AA; 77523 MW; 7FREF7A8F552E3DB CRC64;  
 Query Match 68.8%; Score 33; DB 1; Length 664;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 LVEYFP 8  
 : : : : :  
 DB 248 LVEYFP 253  
 RESULT 12  
 ID ATX1\_MOUSE STANDARD; PRT; 792 AA.  
 AC P54254;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Ataxin-1 (Spinocerebellar ataxia type 1 protein).  
 GN SCAL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6 X CBA; TISSUE=Brain, Thymus, and Retina;  
 RX MEDLINE=96381424; PubMed=8789437;  
 RA Banfi S., Servadio A., Chung M.-Y., Capozzoli F., Duvick L.A.,  
 RA Elde R., Zoghbi H.Y., Orr H.T.;  
 RT "Cloning and developmental expression analysis of the murine homolog  
 RT of the spinocerebellar ataxia type 1 gene (Scal).";  
 RL Hum. Mol. Genet. 5:33-40 (1996).  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. IN BRAIN, THE PATTERN OF  
 CC DISTRIBUTION IS LIMITED TO NEURONS POPULATIONS.  
 CC -1- DEVELOPMENTAL STAGE: TRANSIENT EXPRESSION BURST IN PURKINJE CELLS  
 CC AS THE CEREBELLAR CORTEX BECOMES FUNCTIONAL (POSTNATAL DAY 14),  
 CC AND IN MESENCHYMAL CELLS OF THE DEVELOPING INTERVERTEBRAL DISCS OF  
 CC THE SPINAL COLUMN.  
 CC -1- POLYMORPHISM: THE MURINE POLY-GLN REGION IS VERY LIMITED IN  
 CC COMPARISON TO THE HUMAN SCAL AND IS NOT POLYMORPHIC.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; X83542; CAA58533.1; -.  
 CC MGD; MGI:104783; SCAL.  
 CC InterPro; IPR003652; Atax\_HMG.  
 CC SMART; SM00536; AXH; 1.  
 CC FT DOMAIN 214 217 POLY-PRO.  
 CC SEQUENCE 792 AA; 84052 MW; CA5F59C0013499DB CRC64;

Query Match 68.8%; Score 33; DB 1; Length 792;  
 Best Local Similarity 100.0%; Pred.No. 66;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LVEYFP 8  
 |||||  
 Db 621 LVEYFP 626

RESULT 13  
 ATX1 HUMAN  
 ID ATX1 HUMAN STANDARD; PRT; 816 AA.  
 AC P54253; Q9UJG2;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ataxin-1 (Spinocerebellar ataxia type 1 protein).  
 GN SCAL OR ATX1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum, and Brain;  
 RX MEDLINE=95038938; PubMed=7951322;  
 RA Banfi S., Servadio A., Chung M.-Y., Kwiatkowski T.J. Jr., McCall A.E.,  
 RA Duvick L.A., Shen Y., Roth E.J., Orr H.T., Zoghbi H.Y.;  
 RT "Identification and characterization of the gene causing type 1  
 RT spinocerebellar ataxia";  
 RL Nat. Genet. 7:513-519(1994).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA Tubby B.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS ARE PRODUCED BY  
 CC -!- ALTERNATIVE SPLICING.  
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED THROUGHOUT THE BODY.  
 CC -!- POLYMORPHISM: THE POLY-GLN REGION OF SCAL IS HIGHLY POLYMORPHIC  
 CC (9 TO 39 REPEATS) IN THE NORMAL POPULATION AND IS EXPANDED TO  
 CC ABOUT 40-81 REPEATS IN SCAL PATIENTS. LONGER EXPANSIONS RESULT IN  
 CC EARLIER ONSET AND MORE SEVERE CLINICAL MANIFESTATIONS OF THE  
 CC DISEASE.  
 CC -!- DISEASE: DEFECTS IN SCAL ARE THE CAUSE OF SPINOCEREBELLAR ATAXIA  
 CC TYPE 1; ALSO KNOWN AS OLIVOPONTOCEREBELLAR ATROPHY I (OPCA I).  
 CC SCAL IS AN AUTOSOMAL DOMINANT NEURODEGENERATIVE DISORDER  
 CC CHARACTERIZED BY PROGRESSIVE NEURONAL LOSS IN THE CEREBELLUM,  
 CC BRAIN STEM AND SPINOCEREBELLAR TRACTS. CLINICAL FEATURES ARE  
 CC CEREBELLAR ATAXIA, DYSPHAGIA, OPHTHALMOPARESIS, MUSCLE WASTING AND  
 CC NEUROPATHY. ONSET OF THE DISEASE USUALLY OCCURS IN THE THIRD OR  
 CC FOURTH DECADE OF LIFE AND DEATH OCCURS TEN TO TWENTY YEARS LATER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; X79204; CAA55793.1; -.  
 CC EMBL; AL009031; CAA15622.1; -.  
 CC Genew; HGNC:10548; SCAL.

DR MIM; 601556; -.  
 DR MIM; 164400; -.  
 DR InterPro; IPR003652; Atax\_HMG.  
 DR SMART; SM00536; AXH; 1.  
 CC KW Polymorphism; Triplet repeat expansion; Alternative splicing.  
 CC FT DOMAIN 197 226 POLY-GLN.  
 CC SEQUENCE 816 AA; 87051 MW; D49BA5DB423D0777 CRC64;

Query Match 68.8%; Score 33; DB 1; Length 816;  
 Best Local Similarity 100.0%; Pred.No. 68;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LVEYFP 8  
 |||||  
 Db 646 LVEYFP 651

RESULT 14  
 BIEA RAT  
 ID BIEA RAT STANDARD; PRT; 295 AA.  
 AC P46874;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Biliverdin reductase A precursor (EC 1.3.1.24) (Biliverdin-IX alpha-  
 DE reductase).  
 GN BLVRA OR BLVR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Kidney;  
 RX MEDLINE=92156147; PubMed=1371282;  
 RA Fakhrai H., Maines M.D.;  
 RT "Expression and characterization of a cDNA for rat kidney biliverdin  
 RT reductase. Evidence suggesting the liver and kidney enzymes are the  
 RT same transcript product";  
 RL J. Biol. Chem. 267:4023-4029(1992).  
 RN [2]

RP MUTAGENESIS.  
 RX MEDLINE=94291657; PubMed=8020496;  
 RA McCoubrey W.K. Jr., Maines M.D.;  
 RT "Site-directed mutagenesis of cysteine residues in biliverdin  
 RT reductase. Roles in substrate and cofactor binding";  
 RL Eur. J. Biochem. 222:597-603(1994).  
 CC -!- FUNCTION: CONVERTS BILIVERDIN TO BILIRUBIN; DISPLAYS TWO DISTINCT  
 CC PH OPTIMA USING A DIFFERENT COFACTOR AT EACH PH: NADH AT THE LOWER  
 CC PH 6.7-6.9 RANGE AND NADPH AT PH 8.5-8.7. NADPH, HOWEVER, IS THE  
 CC PROBABLE COFACTOR IN BIOLOGICAL SYSTEMS.  
 CC -!- CATALYTIC ACTIVITY: Bilirubin + NAD(P)(+) = biliverdin + NAD(P)H.  
 CC -!- COFACTOR: BINDS ONE ZINC ION.  
 CC -!- PATHWAY: FINAL STEP IN HEME METABOLISM.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: TO E. COLI YHHX.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; M81681; AAA40830.1; -.  
 CC InterPro; IPR000683; GFO IDH MCA.  
 CC Pfam; PF01408; GFO IDH MCA; 1.  
 CC Oxidoreductase; NAD; NADP; Zinc.  
 CC PROPEP 1 2  
 CC CHAIN 3 295 BILIVERDIN REDUCTASE A.  
 CC DOMAIN 11 16 POLY-VAL.

```

FT METAL          279      279      ZINC (POTENTIAL).
FT METAL          280      280      ZINC (POTENTIAL).
FT METAL          291      291      ZINC (POTENTIAL).
FT METAL          292      292      ZINC (POTENTIAL).
FT MUTAGEN        73       73       C->A: LOSS OF ACTIVITY.
FT MUTAGEN        78       78       C->A: REDUCED ACTIVITY.
FT MUTAGEN        280      280      C->A: REDUCED ACTIVITY.
FT MUTAGEN        291      291      C->A: REDUCED ACTIVITY.
SQ SEQUENCE       295 AA; 33565 MW; 219C8EA96C150588 CRC64;

Query Match      66.7%; Score 32; DB 1; Length 295;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVEYPT 9
Db 94 LVEYPT 100

RESULT 15
ID BIEA_HUMAN STANDARD; PRT; 296 AA.
AC P53004;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Biliverdin reductase A precursor (EC 1.3.1.24) (Biliverdin-IX alpha-
DE reductase).
GN BLVRA OR BLVR OR BVR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96202961; PubMed=8631357;
RA Maines M.D., Polevoda B.V., Huang T.-J., McCoubrey W.K. Jr.;
RT "Human biliverdin IXalpha reductase is a zinc-metalloprotein.
RT Characterization of purified and Escherichia coli expressed
RT enzymes.";
RL Eur. J. Biochem. 235:372-381(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Komuro A., Tobe T., Nakano Y., Yamaguchi T., Tomita M.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-117 FROM N.A.
RA Cordes M., Wollam C., Carter T.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 3-36; 48-74 AND 228-248.
RC TISSUE=Liver;
RX MEDLINE=93143333; PubMed=8424666;
RA Maines M.D., Trakshel G.M.;
RT "Purification and characterization of human biliverdin reductase.";
RL Arch. Biochem. Biophys. 300:320-326(1993).
RN [5]
RP SEQUENCE OF 3-22.
RC TISSUE=Liver;
RX MEDLINE=95014177; PubMed=7929092;
RA Yamaguchi T., Komoda Y., Nakajima H.;
RT "Biliverdin-IX alpha reductase and biliverdin-IX beta reductase from
RT human liver. Purification and characterization.";
RL J. Biol. Chem. 269:24343-24348(1994).
CC
CC -1- FUNCTION: CONVERTS BILIVERDIN TO BILIRUBIN.
CC -1- CATALYTIC ACTIVITY: Bilirubin + NAD(P) (+) = biliverdin + NAD(P)H.
CC -1- COFACTOR: BINDS ONE ZINC ION. HAS DUAL PH/COFACTOR (NADH, NADPH)
CC SPECIFICITY. USES NADH AT THE ACIDIC PH RANGE (6-6.7) AND NADPH AT
CC THE ALKALINE RANGE (8.5-8.7).
CC -1- PATHWAY: FINAL STEP IN HEME METABOLISM.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: LIVER.

```

```

CC -1- SIMILARITY: TO E.COLI YHHX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X93086; CAA63635.1; -.
DR EMBL; U34877; AAC35588.1; -.
DR EMBL; AC005189; AAC35526.1; -.
DR Genew; HGNC:1062; BLVRA.
DR MIM; 109750; -.
DR InterPro; IPR000683; GFO_IDH_MocA.
DR Pfam; PF01408; GFO_IDH_MocA; 1.
KW Oxidoreductase; NAD; NADP; Zinc.
FT PROPEP 1 2
FT CHAIN 3 296 BILIVERDIN REDUCTASE A.
FT DOMAIN 11 16 POLY-VAL.
FT METAL 280 280 ZINC (POTENTIAL).
FT METAL 281 281 ZINC (POTENTIAL).
FT METAL 292 292 ZINC (POTENTIAL).
FT METAL 293 293 ZINC (POTENTIAL).
FT CONFLICT 3 3 A -> T (IN REF. 2).
FT CONFLICT 154 155 SD -> AG (IN REF. 2).
FT CONFLICT 160 160 D -> E (IN REF. 2).
SQ SEQUENCE 296 AA; 33488 MW; 0DFD3B386F4DFC0A CRC64;

Query Match      66.7%; Score 32; DB 1; Length 296;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVEYPT 9
Db 95 LVEYPT 101

Search completed: April 22, 2003, 12:51:48
Job time : 5.5 secs

```

GenCore version 5.1.4 ps\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:47:43 ; Search time 18 Seconds  
(without alignments)  
103.024 Million cell updates/sec

Title: US-09-674-716B-7  
Perfect score: 48  
Sequence: 1 QQLVEYPT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_page:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	77.1	278	10 Q8W454	Q8W454 arabidopsis
2	37	77.1	467	10 Q65550	Q65550 arabidopsis
3	36	75.0	84	8 Q9B313	Q9B313 neoceratodus
4	36	75.0	247	8 Q9MP34	Q9MP34 bothriomyrm
5	36	75.0	352	8 Q9XNV8	Q9XNV8 delomys dor
6	36	75.0	377	8 Q9XNX5	Q9XNX5 scapteromys
7	36	75.0	377	8 Q9XNV0	Q9XNV0 reithrodon
8	36	75.0	379	8 Q9TGL4	Q9TGL4 cervus elisp
9	36	75.0	379	8 Q9TF64	Q9TF64 spermophilu
10	36	75.0	379	8 Q9TF12	Q9TF12 spermophilu
11	36	75.0	379	8 Q9TF11	Q9TF11 spermophilu
12	36	75.0	379	8 Q9TF08	Q9TF08 spermophilu
13	36	75.0	379	8 Q33887	Q33887 akodon toba
14	36	75.0	379	8 Q33940	Q33940 bolomys uri
15	36	75.0	379	8 Q94QC8	Q94QC8 cervus timo
16	36	75.0	380	8 Q9XNX7	Q9XNX7 oxymycterus

17	36	75.0	380	8 Q9XNX6	Q9XNX6 blarinomys
18	36	75.0	380	8 Q9XNX0	Q9XNX0 thomasomys
19	36	75.0	380	8 Q9XNW9	Q9XNW9 thomasomys
20	36	75.0	380	8 Q9XNW8	Q9XNW8 thomasomys
21	36	75.0	380	8 Q9XNW7	Q9XNW7 thomasomys
22	36	75.0	380	8 Q9XNW6	Q9XNW6 thomasomys
23	36	75.0	380	8 Q9XNW3	Q9XNW3 rhipidomys
24	36	75.0	380	8 Q9XNV7	Q9XNV7 delomys sub
25	36	75.0	380	8 Q9TFX5	Q9TFX5 microtyzomy
26	36	75.0	380	8 Q35970	Q35970 thomasomys
27	36	75.0	380	8 Q34853	Q34853 lenoxus api
28	36	75.0	380	8 Q9MNT7	Q9MNT7 andinomys e
29	36	75.0	380	8 Q9MNT5	Q9MNT5 andalgalm
30	36	75.0	380	8 Q9MNT4	Q9MNT4 tapecomys p
31	36	75.0	380	8 Q9MNT3	Q9MNT3 tapecomys p
32	36	75.0	380	8 Q9MNT2	Q9MNT2 eligmodonti
33	36	75.0	380	8 Q9MNT1	Q9MNT1 gramomys gri
34	36	75.0	380	8 Q9MNT0	Q9MNT0 gramomys dom
35	36	75.0	380	8 Q9MNS8	Q9MNS8 calomys cal
36	36	75.0	380	8 Q9MNS7	Q9MNS7 calomys lep
37	36	75.0	380	8 Q9T7L7	Q9T7L7 microtus ri
38	36	75.0	380	8 Q9MHF9	Q9MHF9 microtus lo
39	36	75.0	380	8 Q94WP7	Q94WP7 calomys mus
40	36	75.0	380	8 Q94V01	Q94V01 microtus oa
41	36	75.0	380	8 Q94UZ9	Q94UZ9 microtus gu
42	36	75.0	380	8 Q956Q2	Q956Q2 acornys igni
43	36	75.0	381	8 Q9TGS0	Q9TGS0 abrothrix o
44	36	75.0	381	8 Q9BIW5	Q9BIW5 abrothrix o
45	36	75.0	381	8 Q9BIW4	Q9BIW4 abrothrix s

#### ALIGNMENTS

#### RESULT 1

Q8W454 ID Q8W454 PRELIMINARY; PRT; 278 AA.  
AC Q8W454; (TREMREL. 20, Created)  
DT 01-MAR-2002 (TREMREL. 20, Last sequence update)  
DT 01-MAR-2002 (TREMREL. 21, Last annotation update)  
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)  
DE Putative calmodulin-binding protein.  
GN A74G31000, F6118.90 OR A74G31000.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.,  
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RT "Arabidopsis Full Length cDNA Clones."  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY062856; AAL32934.1; --  
DR EMBL; AY074504; AAL69488.1; --  
SQ SEQUENCE 278 AA; 31406 MW; 398DCFB65C85A836 CRC64;

Query Match 77.1%; Score 37; DB 10; Length 278;

```

Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPPF 8
DB 145 QHLIEYPPF 152
|:|||||
|:|||||

RESULT 2
O65550
ID O65550 PRELIMINARY; PRT; 467 AA.
AC O65550;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Putative calmodulin-binding protein.
GN P6118.90 OR A74G31000.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Benes V., Rechmann S., Borkova D., Ansoerge W., Hoheisel J.,
RA Mewes H.W., Mayer K., Schueller C.;
RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Benes V., Rechmann S., Borkova D., Ansoerge W., Mewes H.W., Lemcke K.,
RA Mayer K.P.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022198; CAA18193.1; -
DR EMBL; AL161578; CAB79818.1; -
SQ SEQUENCE 467 AA; 52831 MW; 4DB1FB318D553B3D CRC64;

Query Match 77.1%; Score 37; DB 10; Length 467;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPPF 8
DB 334 QHLIEYPPF 341
|:|||||
|:|||||

RESULT 3
Q9B313
ID Q9B313 PRELIMINARY; PRT; 84 AA.
AC Q9B313;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome b (fragment).
GN CYTB
OS Neoceratodus forsteri (Australian lungfish).
OC Chordata; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Dipnoi; Ceratodontiformes; Ceratodontidae; Neoceratodus.
OX NCBI_TaxID=7892;
RN [1]
RP SEQUENCE FROM N.A.
RA Frentiu F., Ovenden J.R., Street R.;
RA "Australian lungfish (Neoceratodus forsteri) have low genetic
RT diversity at allozyme and mitochondrial loci: A conservation alert for
RT a living fossil?";
RT Conserv. Genet. 0:0-0(2001).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

```

```

CC -1- COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF344663; AAK29031.1; -
DR InterPro; IPR00179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c_1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
FT NON TER 1
SQ SEQUENCE 84 AA; 9673 MW; 1E7AFAD921592DAC CRC64;

Query Match 75.0%; Score 36; DB 8; Length 84;
Best Local Similarity 87.5%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QQLVEYPPF 9
DB 27 QQLVEYPPF 34
|:|||||
|:|||||

RESULT 4
Q9MP34
ID Q9MP34 PRELIMINARY; PRT; 247 AA.
AC Q9MP34;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome b (fragment).
GN Cytb.
OS Bothriomyx meridionalis.
OC Chordata; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Aculeata;
OC Formicidae; Dolichoderinae; Bothriomyx.
OX NCBI_TaxID=121499;
RN [1]
RP SEQUENCE FROM N.A.
RA Chiotis M., Jermin L.S., Crozier R.H.;
RA "A molecular framework for the phylogeny of the ant subfamily
RT Dolichoderinae.";
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF146714; AAF66714.1; -
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c_1.
DR PROSITE; PS00193; CYTOCHROME_B_QQ; UNKNOWN_1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
FT NON TER 1
SQ SEQUENCE 247 AA; 29263 MW; CC68F8BA0D065B1A CRC64;

Query Match 75.0%; Score 36; DB 8; Length 247;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QQLVEYPPF 9
DB 201 QQLVEYPPF 208
|:|||||
|:|||||

```

RESULT 5  
Q9XNV8  
ID Q9XNV8 PRELIMINARY; PRT; 352 AA.  
AC Q9XNV8; 352 AA.  
DT 01-NOV-1999 (TRENBLrel. 12, Created)  
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE Cytochrome B (Fragment).  
GN CYTB.  
OS Delonix doralis (striped Atlantic forest rat).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
OC Delonix.  
OK NCBI\_TaxID=89119;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Smith M.F., Patton J.L.;  
RT "Phylogenetic relationships and the radiation of sigmodontine rodents in South America: Evidence from cytochrome b.";  
RL J. Mammal. Evol. 6:89-128(1999).  
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS (BY SIMILARITY).  
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN (BY SIMILARITY).  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
DR EMBL; AF108686; AAD45468.1; -.  
DR InterPro; IPR000179; Cyt.b.b6.  
DR Pfam; PF00032; cytochrome\_b\_c1; 1.  
DR Pfam; PF00033; cytochrome\_b\_n; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
DR PROSITE; PS00193; CYTOCHROME\_B\_QO; UNKNOWN 1.  
DR Electron transport; Heme; Mitochondrion; Respiratory chain;  
KW Transmembrane.  
FT NON\_TER 352  
SQ SEQUENCE 352 AA; 39631 MW; AA408DB627140D6C CRC64;  
Query Match 75.0%; Score 36; DB 8; Length 352;  
Best Local Similarity 87.5%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 QLVPEYPT 9  
DB 341 QVPEYPT 348  
RESULT 6  
Q9XNX5  
ID Q9XNX5 PRELIMINARY; PRT; 377 AA.  
AC Q9XNX5; 377 AA.  
DT 01-NOV-1999 (TRENBLrel. 12, Created)  
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE Cytochrome B (Fragment).  
GN CYTB.  
OS Scapteromys tumidus (swamp rat).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
OC Scapteromys.  
OK NCBI\_TaxID=89126;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Smith M.F., Patton J.L.;  
RT "Phylogenetic relationships and the radiation of sigmodontine rodents in South America: Evidence from cytochrome b.";  
RL J. Mammal. Evol. 6:89-128(1999).  
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS (BY SIMILARITY).  
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN (BY SIMILARITY).  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
DR EMBL; AF108669; AAD45451.1; -.  
DR InterPro; IPR000179; Cyt.b.b6.  
DR Pfam; PF00032; cytochrome\_b\_c1; 1.  
DR Pfam; PF00033; cytochrome\_b\_n; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
DR PROSITE; PS00193; CYTOCHROME\_B\_QO; UNKNOWN 1.  
DR Electron transport; Heme; Mitochondrion; Respiratory chain;  
KW Transmembrane.  
FT NON\_TER 377  
SQ SEQUENCE 377 AA; 42392 MW; 692D67AE20B6F3BB CRC64;  
Query Match 75.0%; Score 36; DB 8; Length 377;  
Best Local Similarity 87.5%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 QLVPEYPT 9  
DB 337 QVPEYPT 344  
RESULT 7  
Q9XNV0  
ID Q9XNV0 PRELIMINARY; PRT; 377 AA.  
AC Q9XNV0; 377 AA.  
DT 01-NOV-1999 (TRENBLrel. 12, Created)  
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE Cytochrome B (Fragment).  
GN CYTB.  
OS Reithrodon auritus (bunny rat).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
OC Reithrodon.  
OK NCBI\_TaxID=56234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Smith M.F., Patton J.L.;  
RT "Phylogenetic relationships and the radiation of sigmodontine rodents in South America: Evidence from cytochrome b.";  
RL J. Mammal. Evol. 6:89-128(1999).  
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS (BY SIMILARITY).  
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN (BY SIMILARITY).  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
DR EMBL; AF108694; AAD45476.1; -.  
DR InterPro; IPR000179; Cyt.b.b6.  
DR Pfam; PF00032; cytochrome\_b\_c1; 1.  
DR Pfam; PF00033; cytochrome\_b\_n; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
DR PROSITE; PS00193; CYTOCHROME\_B\_QO; UNKNOWN 1.  
DR Electron transport; Heme; Mitochondrion; Respiratory chain;  
KW Transmembrane.  
FT NON\_TER 377  
SQ SEQUENCE 377 AA; 42533 MW; 618DB37396EE6B97 CRC64;  
Query Match 75.0%; Score 36; DB 8; Length 377;  
Best Local Similarity 87.5%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 QLVPEYPT 9

```

Db 338 QPVEYPPT 345
| |||||
RESULT 8
Q9TGL4 PRELIMINARY; PRT; 379 AA.
AC Q9TGL4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Cervus elaphus (Red deer).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20229580; PubMed=10764539;
RA Kuwayama R, Ozawa T.;
RT "Phylogenetic relationships among european red deer, wapiti, and sika
deer inferred from mitochondrial DNA sequences.";
RL Mol. Phylogenet. Evol. 15:115-123(2000).
CC -|- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -|- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -|- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF157946; AAD50230.1; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c_1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QO; UNKNOWN 1.
DR Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 379 AA; 42871 MW; 36F71191E46D263B CRC64;

Query Match 75.0%; Score 36; DB 8; Length 379;
Best Local Similarity 87.5%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QPVEYPPT 9
Db 341 QPVEYPPT 348
| |||||
RESULT 10
Q9TF12 PRELIMINARY; PRT; 379 AA.
AC Q9TF12;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Sperophilus lateralis trepidus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
OC Sperophilus.
OX NCBI_TaxID=99842;
RN [1]
RP SEQUENCE FROM N.A.
RA Harrison R.G., Sherman P.W., Yensen E., Hoffmann R.S.,
RA Bogdanowicz S.M.;
RT "A molecular phylogeny of ground squirrels and prairie dogs.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -|- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -|- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF157946; AAD50230.1; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c_1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QO; UNKNOWN 1.
DR Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 379 AA; 43067 MW; 7AD39152077DE39A CRC64;

Query Match 75.0%; Score 36; DB 8; Length 379;
Best Local Similarity 87.5%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QPVEYPPT 9
Db 341 QPVEYPPT 348
| |||||
RESULT 9
Q9TF64 PRELIMINARY; PRT; 379 AA.
AC Q9TF64;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Sperophilus lateralis trepidus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
OC Sperophilus.
OX NCBI_TaxID=100390;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S32;

```



```
Query Match 75.0%; Score 36; DB 8; Length 379;
Best Local Similarity 87.5%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLVYEPPT 9
DB 341 QPVEYPT 348

RESULT 11
Q9TF11 PRELIMINARY; PRT; 379 AA.
AC Q9TF11;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Sperophilus madrensis (Sierra Madre ground squirrel).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Sperophilus.
OX NCBI_TaxID=99842;
RN [1]
RP STRAIN=S98;
RC STRAIN=S98;
RA Harrison R.G., Sherman P.W., Yensen E., Hoffmann R.S.,
RA Bogdanowicz S.M.;
RT "A molecular phylogeny of ground squirrels and prairie dogs.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF157947; AAD50231.1; -.
DR InterPro; IPR000179; Cyt b_b6.
DR Pfam; PF00032; cytochrome_b_c1.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QO; UNKNOWN_1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 379 AA; 43071 MW; B3D3852662B28B2B CRC64;

Query Match 75.0%; Score 36; DB 8; Length 379;
Best Local Similarity 87.5%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLVYEPPT 9
DB 341 QPVEYPT 348

RESULT 12
Q9TF08 PRELIMINARY; PRT; 379 AA.
AC Q9TF08;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Sperophilus lateralis trepidus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Sperophilus.
```

```
OX NCBI_TaxID=100390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEY1069;
RA Harrison R.G., Sherman P.W., Yensen E., Hoffmann R.S.,
RA Bogdanowicz S.M.;
RT "A molecular phylogeny of ground squirrels and prairie dogs.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF157950; AAD50234.1; -.
DR InterPro; IPR000179; Cyt b_b6.
DR Pfam; PF00032; cytochrome_b_c1.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QO; UNKNOWN_1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 379 AA; 42994 MW; 767009AD223C18AD CRC64;

Query Match 75.0%; Score 36; DB 8; Length 379;
Best Local Similarity 87.5%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLVYEPPT 9
DB 341 QPVEYPT 348

RESULT 13
Q33887 PRELIMINARY; PRT; 379 AA.
AC Q33887;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Akodon toba (Chaco grass mouse).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC Akodon.
OX NCBI_TaxID=29101;
RN [1]
RP SEQUENCE OF 1-267 FROM N.A.
RC STRAIN=UMMZ CATALOG # 133965; TISSUE=LIVER;
RA Smith M.F., Patton J.L.;
RT "Diversification of South American murid rodents: Evidence from
RT mitochondrial DNA sequence data for the akodontine tribe.";
RL Biol. J. Linn. Soc. 50:149-177(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=UMMZ CATALOG # 133965; TISSUE=LIVER;
RA Smith M.F., Patton J.L.;
RT "Phylogenetic relationships and the radiation of sigmodontine rodents
RT in South America: Evidence from cytochrome b.";
RL J. Mammal. Evol. 6:89-128(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UMMZ CATALOG # 133965; TISSUE=LIVER;
RA Smith M.F.;
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
```

CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).  
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN (BY SIMILARITY).  
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
DR EMBL; U03527; AAD12554.2; -.  
DR InterPro; IPR000179; Cyt\_b\_b6.  
DR Pfam; PF00032; cytochrome\_b\_c1.  
DR Pfam; PF00033; cytochrome\_b\_n; 1.  
DR PROSITE; PS00192; CYTOCHROME B HEME; 1.  
DR PROSITE; PS00193; CYTOCHROME B\_QO; UNKNOWN 1.  
KW Electron transport; Heme; Mitochondrion; Respiratory chain;  
KW Transmembrane.  
SQ SEQUENCE 379 AA; 42689 MW; 8B946C25430PDCD0 CRC64;  
  
Query Match 75.0%; Score 36; DB 8; Length 379;  
Best Local Similarity 87.5%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 QLVVEPPT 9  
Db 341 QPVEYPPT 348  
| | | | |  
  
RESULT 14  
Q33940 PRELIMINARY; PRT; 379 AA.  
AC Q33940;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Cytochrome b.  
DN CYTB.  
OS Bolomys urichi.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
OC Bolomys.  
OX NCBI\_TaxID=100644;  
RN [1]\_TaxID=100644;  
RP SEQUENCE OF 1-267 FROM N.A.  
RC STRAIN-USNM CATALOG # 560661; TISSUE=LIVER;  
RA Smith M.F., Patton J.L.;  
RT "Diversification of South American murid rodents: Evidence from  
RT mitochondrial DNA sequence data for the akodontine tribe.";  
RL Biol. J. Linn. Soc. 50:149-177(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-USNM CATALOG # 560661; TISSUE=LIVER;  
RA Smith M.F., Patton J.L.;  
RT "Phylogenetic relationships and the radiation of sigmodontine rodents  
RT in South America: Evidence from cytochrome b.";  
RL J. Mammal. Evol. 6:89-128(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-USNM CATALOG # 560661; TISSUE=LIVER;  
RA Smith M.F.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE  
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A  
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL  
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).  
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN (BY SIMILARITY).  
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
DR EMBL; U03549; AAD12575.2; -.  
DR InterPro; IPR000179; Cyt\_b\_b6.  
DR Pfam; PF00032; cytochrome\_b\_c1.  
DR Pfam; PF00033; cytochrome\_b\_n; 1.  
DR PROSITE; PS00192; CYTOCHROME B\_HEME; 1.  
DR PROSITE; PS00193; CYTOCHROME B\_QO; UNKNOWN 1.  
KW Electron transport; Heme; Mitochondrion; Respiratory chain;  
KW Transmembrane.  
SQ SEQUENCE 379 AA; 42689 MW; 8B946C25430PDCD0 CRC64;

DR PROSITE; PS00193; CYTOCHROME B\_QO; UNKNOWN 1.  
KW Electron transport; Heme; Mitochondrion; Respiratory chain;  
KW Transmembrane.  
SQ SEQUENCE 379 AA; 42766 MW; 8B74E09004D8B663 CRC64;  
  
Query Match 75.0%; Score 36; DB 8; Length 379;  
Best Local Similarity 87.5%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 QLVVEPPT 9  
Db 341 QPVEYPPT 348  
| | | | |  
  
RESULT 15  
Q94QC8 PRELIMINARY; PRT; 379 AA.  
AC Q94QC8;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Cytochrome b.  
DN CYTB.  
OS Cervus timorensis macassanicus.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;  
OC Cervidae; Cervinae; Cervus.  
OX NCBI\_TaxID=173988;  
RN [1]\_TaxID=173988;  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPONGIOSA;  
RA Lutz C.J., Kuehn R., Schroeder W., Rottmann O.;  
RT "Phylogeny of Cervidae based on mitochondrial genes";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE  
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A  
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL  
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).  
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN (BY SIMILARITY).  
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
DR EMBL; AF423200; AAL17842.1; -.  
DR InterPro; IPR000179; Cyt\_b\_b6.  
DR Pfam; PF00032; cytochrome\_b\_c1.  
DR Pfam; PF00033; cytochrome\_b\_n; 1.  
DR PROSITE; PS00192; CYTOCHROME B\_HEME; UNKNOWN 1.  
DR PROSITE; PS00193; CYTOCHROME B\_QO; UNKNOWN 1.  
KW Electron transport; Heme; Mitochondrion; Respiratory chain;  
KW Transmembrane.  
SQ SEQUENCE 379 AA; 42915 MW; F2337185154459B3 CRC64;  
  
Query Match 75.0%; Score 36; DB 8; Length 379;  
Best Local Similarity 87.5%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 QLVVEPPT 9  
Db 341 QPVEYPPT 348  
| | | | |  
  
Search completed: April 22, 2003, 12:53:49  
Job time : 19 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:45:19 ; Search time 22.8214 Seconds  
(without alignments)  
52.550 Million cell updates/sec

Title: US-09-674-716B-7

Perfect score: 48

Sequence: 1 QQLVEYPT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	48	100.0	9	21 AAY32256
2	48	100.0	116	21 AAY32262
3	48	100.0	145	21 AAY32261
4	45	93.8	19	21 AAY70804
5	45	93.8	122	21 AAY70790
6	45	93.8	131	12 AAR12232
7	45	93.8	132	12 AAR12354
8	44	91.7	9	19 AAW39823
9	44	91.7	113	19 AAW39882
10	44	91.7	113	19 AAW39804

11	37	77.1	100	22	AAE06969	Mouse germline kap
12	36	75.0	9	19	AAW39817	Light chain CDR3 o
13	36	75.0	113	19	AAW39801	Variable domain of
14	36	75.0	113	19	AAW39803	Variable domain of
15	36	75.0	241	18	AAW24063	Human WSX receptor
16	36	75.0	241	23	AAU09048	Insulin/insulin-li
17	36	75.0	251	23	ABP45299	Human BlyS binding
18	36	75.0	253	23	ABP44925	Human BlyS binding
19	36	75.0	274	19	AAW39899	Single chain Fv re
20	36	75.0	392	21	AAE10863	S11-VEGF2 construc
21	36	75.0	510	21	AAE10864	S11-VEGF2 constr
22	35	72.9	112	21	AAW39813	2G3 hybridoma VL d
23	35	72.9	126	20	AAW36951	Protein involved i
24	35	72.9	533	21	AAW36951	Arabidopsis thalia
25	35	72.9	536	21	AAW36951	Arabidopsis thalia
26	35	72.9	622	21	AAW36951	Arabidopsis thalia
27	34	70.8	9	21	AAW36951	Murine 13H10 light
28	34	70.8	113	21	AAW36951	Murine 13H10 light
29	34	70.8	146	18	AAW28154	Amino acid sequenc
30	34	70.8	585	23	ABP40069	Staphylococcus epi
31	34	70.8	604	22	AAW36951	S. epidermidis ope
32	33	68.8	9	14	AAW36951	C242:11 MAB kappa
33	33	68.8	82	22	AAW36951	Peptide #4456 enco
34	33	68.8	107	22	AAW36951	Human anti-Rh(D) c
35	33	68.8	108	14	AAW36951	BW 835 VK. Synthe
36	33	68.8	113	15	AAW36951	Sequence of the ma
37	33	68.8	113	21	AAW36951	260F9 hybridoma VL
38	33	68.8	133	14	AAW36951	gLI variable domai
39	33	68.8	133	14	AAW36951	CTM01 VL. Synthe
40	33	68.8	133	15	AAW36951	Sequence of the li
41	33	68.8	133	15	AAW36951	Sequence of the li
42	33	68.8	133	18	AAW36951	CDR-grafted humani
43	33	68.8	133	21	AAW36951	MAB CT-M-01 light
44	33	68.8	133	21	AAW36951	gLI variable domai
45	33	68.8	139	22	AAO00042	Human polypeptide

## ALIGNMENTS

RESULT 1

AAW32256

ID AAY32256 standard; Peptide; 9 AA.

XX AAY32256;

XX 15-FEB-2000 (first entry)

DT Light chain CDR L3 of mouse anti-CD23 MAB C11.

DE CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse;  
KW monoclonal antibody; chimeric antibody; humanised antibody;  
KW complementarity determining region; CDR; autoimmune disease;  
KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
KW urticaria; nephrotic syndrome; glomerulonephritis;  
KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
KW graft-versus-host disease; COPD; bronchitis; diabetes;  
KW B-cell malignancy; therapy.

OS Mus musculus.

PN WO958679-A1.

PD 18-NOV-1999.

XX 07-MAY-1999; 99WO-GB01434.

XX 09-MAY-1998; 98GB-0009839.

XX (GLAXO) GLAXO GROUP LTD.

PI Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 XX WPI; 2000-053101/04.  
 DR N-PSDB; AA324741.  
 XX  
 XX Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 PT  
 XX Claim 1; Page 40; 81pp; English.  
 PS  
 XX This sequence represents complementarity determining region 3  
 CC (CDR L3) of the light chain of murine anti-CD23 (FCERII) monoclonal  
 CC antibody C11 (see also AAY32262). The invention provides altered  
 CC antibodies, such as chimeric or humanised antibodies, which comprise  
 CC sufficient of the amino acid sequences of C11 light and heavy chain  
 CC CDRs (see AAY32254-59) to render them capable of binding to the CD23  
 CC type II molecule expressed on haematopoietic cells. The antibodies  
 CC are used to block soluble CD23 formation for treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX  
 XX Sequence 9 AA;  
 SQ  
 Query Match 100.0%; Score 48; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 QQLVEVPFT 9  
 Db 1 QQLVEVPFT 9  
 RESULT 2  
 AAY32262  
 ID AAY32262 standard; Protein; 116 AA.  
 XX  
 AC AAY32262;  
 XX  
 XX 15-FEB-2000 (first entry)  
 DT  
 XX Humanised anti-CD23 Mab C11 light chain variable region.  
 XX  
 XX CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse; human;  
 KW monoclonal antibody; chimeric antibody; humanised antibody;  
 KW complementarity determining region; CDR; autoimmune disease;  
 KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
 KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
 KW urticaria; nephrotic syndrome; glomerulonephritis;  
 KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
 KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
 KW graft-versus-host disease; COPD; bronchitis; diabetes;  
 KW B-cell malignancy; therapy.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..23  
 FT /note= "framework region 1"  
 FT Region 24..39  
 FT /note= "CDR 1"  
 FT Region 40..54  
 FT /note= "framework region 2"  
 FT Region 55..61  
 FT /note= "CDR 2"

FT Region 62..93  
 FT /note= "framework region 3"  
 FT Region 94..102  
 FT /note= "CDR 3"  
 FT Region 103..113  
 FT /note= "framework region 4"  
 XX  
 XX WO958679-A1.  
 XX  
 XX 18-NOV-1999.  
 XX  
 XX 07-MAY-1999; 99WO-GB01434.  
 XX  
 XX 09-MAY-1998; 98GB-0009839.  
 XX (GLAX) GLAXO GROUP LTD.  
 XX  
 XX Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 XX WPI; 2000-053101/04.  
 DR N-PSDB; AA324741.  
 DR  
 XX Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 PT  
 XX Claim 9; Fig 3; 81pp; English.  
 PS  
 XX This sequence represents the light chain variable region (VL) of  
 CC humanised anti-CD23 (FCERII) monoclonal antibody C11, composed of  
 CC a human framework (H5IGKVII) and the light chain complementarity  
 CC determining regions (see AAY32254-56) of murine antibody C11. The  
 CC DNA was constructed by splice overlap PCR. The invention provides  
 CC altered antibodies, such as chimeric or humanised antibodies, which  
 CC comprise sufficient of the amino acid sequences of the C11 light  
 CC and heavy chain complementarity determining regions to render them  
 CC capable of binding to the CD23 type II molecule expressed on  
 CC haematopoietic cells. The antibodies are used to block soluble  
 CC CD23 formation in human therapy, for the treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX  
 XX Sequence 116 AA;  
 SQ  
 Query Match 100.0%; Score 48; DB 21; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 0.053;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 QQLVEVPFT 9  
 Db 94 QQLVEVPFT 102  
 RESULT 3  
 AAY32261  
 ID AAY32261 standard; Protein; 145 AA.  
 XX  
 AC AAY32261;  
 XX  
 XX 15-FEB-2000 (first entry)  
 DT  
 XX Mouse anti-CD23 Mab C11 light chain variable region.  
 DE  
 XX CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse;  
 KW monoclonal antibody; chimeric antibody; humanised antibody;  
 KW complementarity determining region; CDR; autoimmune disease;  
 KW

KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
 KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
 KW urticaria; nephrotic syndrome; glomerulonephritis;  
 KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
 KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
 KW graft-versus-host disease; COPD; bronchitis; diabetes;  
 KW B-cell malignancy; therapy.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 55..70  
 FT /note= "CDR L1"  
 FT Region 83..92  
 FT /note= "CDR L2"  
 FT Region 125..134  
 FT /note= "CDR L3"  
 FT  
 PN WO958679-A1.  
 XX  
 XX 18-NOV-1999.  
 PD  
 XX  
 XX 07-MAY-1999; 99WO-GB01434.  
 PF  
 XX  
 XX 09-MAY-1999; 98GB-0009839.  
 PR  
 XX  
 XX (GLAX ) GLAXO GROUP LTD.  
 PA  
 XX  
 XX Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 PI WPI; 2000-053101/04.  
 XX  
 XX N-PSDB; AA234746.  
 DR  
 XX  
 XX Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 PT  
 XX  
 PS Claim 8; Fig 2; 81pp; English.  
 XX  
 CC This sequence represents the light chain variable region (VL) of  
 CC murine anti-CD23 (FCBR1) monoclonal antibody C11. The invention  
 CC provides altered antibodies, such as chimeric or humanised  
 CC antibodies (see AAY32262 and AAY32263), which comprise sufficient of  
 CC the amino acid sequences of the C11 light and heavy chain  
 CC complementarity determining regions (see AAY32254-59) to render them  
 CC capable of binding to the CD23 type II molecule expressed on  
 CC haematopoietic cells. The antibodies are used to block soluble  
 CC CD23 formation in human therapy, for the treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX  
 SQ Sequence 145 AA;  
 Query Match 100.0%; Score 48; DB 21; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 0.067;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QQLVEYPFT 9  
 DB 125 QQLVEYPFT 133  
 RESULT 4  
 AAY70804  
 ID AAY70804 standard; peptide; 19 AA.  
 XX

AC AAY70804;  
 XX  
 XX 31-JUL-2000 (first entry)  
 DT  
 DE Murine anti-PAb-421 IDI-1 mAb light chain CDR based peptide IDI-1 L3.  
 XX  
 XX Murine; p53 protein; monoclonal antibody; mAb; PAb-421; IDI-1 L3;  
 KW light chain variable region; VL; complementarity determining region; CDR;  
 KW dermatological; immunosuppressive; antiinflammatory; autoimmune response;  
 KW SLE; systemic lupus erythematosus; diagnosis; treatment; autoantigen;  
 KW DNA-binding domain; anti-idiotypic antibody.  
 XX  
 OS Mus sp.  
 XX  
 XX WO200023082-A1.  
 PN  
 XX 27-APR-2000.  
 PD  
 XX 19-OCT-1999; 99WO-US24443.  
 PF  
 XX 19-OCT-1998; 98US-0104816.  
 PR  
 XX (YEDA ) YEDA RES & DEV CO LTD.  
 PA  
 XX Cohen IR, Rotter V, Erez-Alon N, Herkel J;  
 PI WPI; 2000-339512/29.  
 XX  
 XX Treatment of systemic lupus erythematosus by down-regulating the  
 FT autoimmune response to the C-terminal DNA-binding domain of the p53  
 FT protein by an active compound comprising of antibodies to p53 or  
 FT fragments of p53 -  
 FT  
 XX Claim 78; Fig 10; 87pp; English.  
 PS  
 XX The patent discloses a method for the treatment of systemic lupus  
 CC erythematosus (SLE) by down-regulating the autoimmune response to the  
 CC C-terminal DNA-binding domain of p53 protein by an active compound.  
 CC The present sequence is a IDI-1 L3 peptide which comprises the  
 CC complementarity determining region (CDR) of the light chain of  
 CC IDI-1 monoclonal antibody (mAb). The IDI-1 mAb is an anti-idiotypic  
 CC antibody/Ab2 mAb specific for PAb-421 which is an Ab1 mAb specific  
 CC to the C-terminal DNA-binding domain of murine p53 protein.  
 CC The peptide corresponds to residues 92-110 of IDI-1 light chain  
 CC variable region. It is an example of the active compound useful in  
 CC the diagnosis, prevention and treatment of SLE in humans.  
 XX  
 XX Sequence 19 AA;  
 Query Match 93.8%; Score 45; DB 21; Length 19;  
 Best Local Similarity 88.9%; Pred. No. 0.031;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QQLVEYPFT 9  
 DB 6 QQLVEYPFT 14  
 RESULT 5  
 AAY70790  
 ID AAY70790 standard; Protein; 122 AA.  
 XX  
 XX AC AAY70790;  
 XX  
 XX 31-JUL-2000 (first entry)  
 DT  
 XX Murine anti-PAB-421 IDI-1 mAb light chain variable region.  
 DE  
 XX Murine; p53 protein; PAb-421; monoclonal antibody; mAb; IDI-1;  
 KW anti-idiotypic antibody; DNA-binding domain;  
 KW dermatological; immunosuppressive; antiinflammatory; autoimmune response;  
 KW SLE; systemic lupus erythematosus; diagnosis; treatment; autoantigen;  
 KW light chain variable region; VL; complementarity determining region; CDR.

```

XX OS Mus sp.
XX PF
XX FT Location/Qualifiers
XX FT 27..42
XX FT /label= CDR
XX FT /note= "Complementarity determining region"
XX FT 58..64
XX FT /label= CDR
XX FT /note= "Complementarity determining region"
XX FT 97..105
XX FT /label= CDR
XX FT /note= "Complementarity determining region"
XX PN WO200023082-A1.
XX PD
XX PD 27-APR-2000.
XX PF 19-OCT-1999; 99WO-US24443.
XX PR 19-OCT-1998; 98US-0104816.
XX PA (YEDA ) YEDA RES & DEV CO LTD.
XX PI Cohen IR, Rotter V, Brez-Alon N, Herkel J;
XX DR WPI; 2000-339512/29.
XX PT Treatment of systemic lupus erythematosus by down-regulating the
XX PT autoimmunity response to the C-terminal DNA-binding domain of the p53
XX PT protein by an active compound comprising of antibodies to p53 or
XX PT fragments of p53 -
XX PS Claim 78; Fig 9; 87pp; English.
XX CC The patent discloses a method for the treatment of systemic lupus
XX CC erythematosus (SLE) by down-regulating the autoimmunity response to the
XX CC C-terminal DNA-binding domain of p53 protein by an active compound.
XX CC The present sequence is a light chain variable region of ID1-1 an
XX CC anti-idiotypic antibody/Ab2 monoclonal antibody (mAb) specific for
XX CC PAB-421 which is an Ab1 mAb specific to the C-terminal DNA-binding domain
XX CC of murine p53 protein. The Ab1 and Ab2 mAbs and peptides based on
XX CC complementarity determining regions of light and heavy chain variable
XX CC regions of these antibodies, are examples of active compounds useful in
XX CC the diagnosis, prevention and treatment of SLE in humans.
XX SQ Sequence 122 AA;
XX Query Match 93.8%; Score 45; DB 21; Length 122;
XX Best Local Similarity 88.9%; Pred. No. 0.22;
XX Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QQLVEYPT 9
DB 97 QQLVEYPT 105
RESULT 6
AAR12232
ID AAR12232 standard; Protein; 131 AA.
XX AC
XX AC AAR12232;
XX DT 19-AUG-1991 (first entry)
XX DE Mouse MAb 2E12 L chain V region.
XX KM HTV-1; chimera.
XX OS Mus sp.
XX PN WO9107494-A.
XX CC

```

```

PD 30-MAY-1991.
XX 13-NOV-1990; 90WO-US06627.
XX* 13-NOV-1989; 89US-0433703.
XX (XOMA-) XOMA CORP.
XX Better MD, Horwitz AH, Ghosh-Dastidar P, Robinson RR;
XX WPI; 1991-178106/24.
XX N-PSDB; AAQ12012.
XX New chimeric mouse human antibodies - used in treatment, diagnosis
XX PT and prophylaxis of HIV infections.
XX PS Disclosure; Fig 1; 108pp; English.
XX CC The mouse VL gene product may be used to produce chimeric mouse-
XX CC human Abs against HIV-1 comprising human Ig constant regions and
XX CC murine variable regions. These novel sequence are useful in
XX CC treatment, diagnosis and prophylaxis of HIV infections, and may be
XX CC produced by a bacterial, yeast or mammalian expression system.
XX SQ Sequence 131 AA;
XX Query Match 93.8%; Score 45; DB 12; Length 131;
XX Best Local Similarity 88.9%; Pred. No. 0.24;
XX Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QQLVEYPT 9
DB 114 QQLVEYPT 122
RESULT 7
AAR12354
ID AAR12354 standard; Protein; 132 AA.
XX AC
XX AC AAR12354;
XX DT 15-AUG-1991 (first entry)
XX DE Light (kappa) chain variable region of murine 2E12
XX DE immunoglobulin.
XX KW Chimeric antibodies; immunoconjugates; HIV; AIDS.
XX OS Mus musculus.
XX PN WO9107493-A.
XX PD 30-MAY-1991.
XX 13-NOV-1990; 90WO-US06615.
XX 13-NOV-1989; 89US-0433730.
XX (XOMA-) XOMA CORP.
XX (GREC ) GREEN CROSS CORP.
XX Better MD, Horwitz AH, Ghosh-Dastidar P, Robinson RR;
XX WPI; 1991-178044/24.
XX N-PSDB; AAQ12056.
XX New chimeric mouse-human antibodies - used to detect, kill and
XX PT remove HIV-1 antigen from sample
XX PS Disclosure; fig 1; 107pp; English.
XX CC This is the light (kappa) - chain variable (V) region of a mouse
XX CC monoclonal antibody (Mab), 2E12, and is specific for an HIV-1

```

CC viral antigen. It is used in the construction of a chimeric Mab  
 CC comprising heavy and light chains having murine V regions and human  
 CC C regions. The chimeric Mabs are more effective than murine Mab  
 CC 2E12 since they have an increased compatibility in humans. The  
 CC heavy and light chain V-regions are joined by manipulating their  
 CC respective joining (J) regions, to generate restriction enzyme  
 CC recognition sites. The chimeric Mabs can be used as immuno-  
 CC conjugates, in association with e.g. toxins for HIV treatment.  
 CC They can also be used in diagnosis of HIV.  
 CC See also AAK12057-63.  
 SQ Sequence 132 AA;

Query Match 93.8%; Score 45; DB 12; Length 132;  
 Best Local Similarity 88.9%; Pred. No. 0.24; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 1;

QY 1 QQLVEYPFT 9  
 |||||:  
 Db 114 QQLVEYPFT 122

RESULT 8  
 AAW39823  
 ID AAW39823 standard; peptide; 9 AA.

XX AAW39823;

XX 16-JUN-1998 (first entry)

XX Light chain CDR3 of catalytic antibody 2A10.

XX Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.

XX Mus sp.

XX WO9749800-A1.

XX 31-DEC-1997.

XX 25-JUN-1997; 97WO-US10965.

XX 25-JUN-1996; 96US-0672345.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Landry DW;

XX WPI; 1998-077166/07.

XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding

XX Claim 15; Page 82; 147pp; English.

XX AAW39821-23 represent the sequences of the light chain complementarity  
 CC determining regions (CDRs) of the catalytic antibody 2A10, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 2A10 antibody was  
 CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
 CC monoester transition state analogue. Antibody 2A10 has a per minute Kcat  
 CC of 0.011. The antibodies reduce the concentration of cocaine in a  
 CC subject, and are used particularly for the treatment of an overdose. They  
 CC are also used for treating addiction (by reducing the in vivo  
 CC concentration that can be achieved).

XX Sequence 9 AA;

Query Match 91.7%; Score 44; DB 19; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 7.8e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
 |||||:  
 Db 1 QQFVEYPFT 9

RESULT 9  
 AAW39882

ID AAW39882 standard; Protein; 113 AA.

XX AAW39882;

XX 16-JUN-1998 (first entry)

XX Light chain of the catalytic antibody 2A10.

XX Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.

XX Mus sp.

XX WO9749800-A1.

XX 31-DEC-1997.

XX 25-JUN-1997; 97WO-US10965.

XX 25-JUN-1996; 96US-0672345.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Landry DW;

XX WPI; 1998-077166/07.

XX N-PSDB; AAV09789.

XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding

XX Disclosure; Fig 21; 147pp; English.

XX The present sequence represents the light chain of a catalytic antibody  
 CC which is capable of degrading cocaine. A series of cocaine transition  
 CC state analogues (TSAs) were prepared and used to immunise mice for  
 CC production of hybridomas. Catalytic antibodies were identified by their  
 CC capacity to release 3H-benzoic acid from 3H-phenyl cocaine. The  
 CC antibodies reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used for  
 CC treating addiction (by reducing the in vivo concentration that can be  
 CC achieved).

XX Sequence 113 AA;

Query Match 91.7%; Score 44; DB 19; Length 113;  
 Best Local Similarity 88.9%; Pred. No. 0.33;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
 |||||:  
 Db 94 QQFVEYPFT 102

RESULT 10  
 AAW39804

ID AAW39804 standard; protein; 113 AA.

XX AAW39804;  
 XX 16-JUN-1998 (first entry)  
 XX Variable domain of the Kappa light chain of catalytic antibody 2A10.  
 XX  
 XX Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX  
 XX Mus sp.  
 XX WO9749800-A1.  
 XX  
 XX 31-DEC-1997.  
 XX  
 XX 25-JUN-1997; 97WO-US10965.  
 XX  
 XX 25-JUN-1996; 96US-0672345.  
 XX  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 XX Landry DW;  
 XX WPI; 1998-077166/07.  
 XX  
 XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX  
 XX Claim 16; Pages 73-74; 147pp; English.  
 XX  
 XX AAW39801-05 represent the amino acid sequences of the variable domain  
 CC of the Kappa light chain of catalytic antibodies which are able to  
 CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
 CC were prepared and used to immunise mice for production of hybridomas.  
 CC Catalytic antibodies were identified by their capacity to release  
 CC 3H-benzoic acid from 3H-phenyl cocaine. The 2A10 antibody (AAW39809  
 CC represents the heavy chain) was identified using TSA1, which is an  
 CC immunogenic conjugate of a phosphate monoester transition state  
 CC analogue. Antibody 2A10 has a per minute Kcat of 0.011. The antibodies  
 CC reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used  
 CC for treating addiction (by reducing the in vivo concentration that can  
 CC be achieved).  
 XX  
 XX Sequence 113 AA;  
 Query Match 91.7%; Score 44; DB 19; Length 113;  
 Best Local Similarity 88.9%; Pred. No. 0.33; Indels 0; Gaps 0;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QQLVEYPT 9  
 Db 94 QQLVEYPT 102  
 RESULT 11  
 AAE06969  
 ID AAE06969 standard; Protein; 100 AA.  
 XX  
 XX AAE06969;  
 XX  
 XX 16-OCT-2001 (first entry)  
 XX  
 XX Mouse germline kappa light chain variable (VK) region, 167/24.  
 XX  
 XX Mouse; humanised antibody; CC-chemokine receptor 2; CCR2; nephroretropic;  
 KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;

KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
 KW multiple sclerosis; atherosclerosis; stenosis; asthma;  
 KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
 KW AIDS; inflammatory glomerulopathy; vascular intervention;  
 KW neointimal hyperplasia; VK; kappa light chain variable region.  
 XX  
 XX Mus sp.  
 XX WO200157226-A1.  
 XX  
 XX 09-AUG-2001.  
 XX  
 XX 02-FEB-2001; 2001WO-US03537.  
 XX  
 XX 03-FEB-2000; 2000US-0497625.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
 XX WPI; 2001-488888/53.  
 XX  
 XX Humanized immunoglobulin for treating a CC-chemokine receptor  
 PT 2-mediated disorder in a patient, comprises a binding specificity for  
 PT CCR2, and a non-human antigen binding region and human immunoglobulin  
 PT  
 XX  
 XX Disclosure; Page 151; 183pp; English.  
 XX  
 XX The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherosclerosis and atrophy of diagnosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR-2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is mouse germline kappa light chain variable  
 CC (VK) region, 167/24.  
 XX  
 XX Sequence 100 AA;  
 Query Match 77.1%; Score 37; DB 22; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 7.3;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QQLVEYPT 7  
 Db 94 QQLVEYPT 100  
 RESULT 12  
 AAW39817  
 ID AAW39817 standard; peptide; 9 AA.  
 XX  
 XX AAW39817;  
 AC  
 XX 16-JUN-1998 (first entry)  
 XX



DE Light chain CDR3 of catalytic antibody 3B9.  
 XX Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX Mus sp.  
 OS WO9749800-A1.  
 PN 31-DEC-1997.  
 XX  
 PD 25-JUN-1997; 97WO-US10965.  
 XX  
 PF 25-JUN-1996; 96US-0672345.  
 XX  
 PR (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PA Landry DW;  
 XX  
 PI WPI; 1998-077166/07.  
 XX  
 DR New catalytic antibodies able to decompose cocaine, single-chain  
 XX analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 PT  
 XX Claim 11; Page 81; 147pp; English.  
 PS AAW39815-17 represent the sequences of the light chain complementarity  
 XX determining regions (CDRs) of the catalytic antibody 3B9, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 3B9 antibody was  
 CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
 CC monoester transition state analogue. Antibody 3B9 has a per minute Kcat  
 CC of 0.11. The antibodies reduce the concentration of cocaine in a subject,  
 CC and are used particularly for the treatment of an overdose. They are also  
 CC used for treating addiction (by reducing the in vivo concentration that  
 CC can be achieved).  
 XX  
 SQ Sequence 9 AA;  
 Query Match 75.0%; Score 36; DB 19; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 QQLVEYPPT 9  
 Db 1 QHFVDYPPT 9  
 RESULT 13  
 AAW39801  
 ID AAW39801 standard; protein; 113 AA.  
 XX  
 AC AAW39801;  
 XX  
 DT 16-JUN-1998 (first entry)  
 XX  
 DE Variable domain of the Kappa light chain of catalytic antibody 3B9.  
 XX  
 KW Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9749800-A1.  
 XX  
 PD 31-DEC-1997.  
 XX  
 PF 25-JUN-1997; 97WO-US10965.  
 XX  
 PR 25-JUN-1996; 96US-0672345.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Landry DW;  
 XX

PD 31-DEC-1997.  
 XX  
 PF 25-JUN-1997; 97WO-US10965.  
 XX  
 PR 25-JUN-1996; 96US-0672345.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Landry DW;  
 XX  
 DR WPI; 1998-077166/07.  
 XX  
 DR P-PSDB; AAV09791.  
 XX  
 PT New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 PT  
 XX Claim 12; Pages 71-72; 147pp; English.  
 PS AAW39801-05 represent the amino acid sequences of the variable domain  
 XX of the Kappa light chain of catalytic antibodies which are able to  
 CC degrade cocaine. A series of cocaine transition state analogues (TSAs)  
 CC were prepared and used to immunise mice for production of hybridomas.  
 CC Catalytic antibodies were identified by their capacity to release  
 CC 3H-benzoic acid from 3H-phenyl cocaine. The 3B9 antibody (AAW39806  
 CC represents the heavy chain) was identified using TSA1, which is an  
 CC immunogenic conjugate of a phosphate monoester transition state  
 CC analogue. Antibody 3B9 has a per minute Kcat of 0.11. The antibodies  
 CC reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used  
 CC for treating addiction (by reducing the in vivo concentration that can  
 CC be achieved).  
 XX  
 SQ Sequence 113 AA;  
 Query Match 75.0%; Score 36; DB 19; Length 113;  
 Best Local Similarity 66.7%; Pred. No. 13;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 QQLVEYPPT 9  
 Db 94 QHFVDYPPT 102  
 RESULT 14  
 AAW39803  
 ID AAW39803 standard; protein; 113 AA.  
 XX  
 AC AAW39803;  
 XX  
 DT 16-JUN-1998 (first entry)  
 XX  
 DE Variable domain of the Kappa light chain of catalytic antibody 12H1.  
 XX  
 KW Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9749800-A1.  
 XX  
 PD 31-DEC-1997.  
 XX  
 PF 25-JUN-1997; 97WO-US10965.  
 XX  
 PR 25-JUN-1996; 96US-0672345.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Landry DW;  
 XX

```

XX DR WPI; 1998-077166/07.
XX DR P-PSDB; AAV09802.
XX PT New catalytic antibodies able to decompose cocaine, single-chain
XX PT analogues - used to treat cocaine overdose and addiction, required
XX PT in far smaller doses than antibodies that antagonise cocaine by
XX PT simply binding
XX PS Claim 18; Page 73; 147pp; English.
XX CC AAW39801-05 represent the amino acid sequences of the variable domain
XX CC of the Kappa light chain of catalytic antibodies which are able to
XX CC degrade cocaine. A series of cocaine transition state analogues (TSAs)
XX CC were prepared and used to immunise mice for production of hybridomas.
XX CC Catalytic antibodies were identified by their capacity to release
XX CC 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody (AAW39808
XX CC represents the heavy chain) was identified using TSAs, and has a per
XX CC minute Kcat of 0.016. The antibodies reduce the concentration of cocaine
XX CC in a subject, and are used particularly for the treatment of an
XX CC overdose. They are also used for treating addiction (by reducing the in
XX CC vivo concentration that can be achieved).
XX SQ Sequence 113 AA;
    Query Match 75.0%; Score 36; DB 19; Length 113;
    Best Local Similarity 66.7%; Pred. No. 13;
    Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 QQLVVEYPPT 9
Db 94 QHFVDYPT 102
    | | | | |
    | | | | |

RESULT 15
AAW24063
ID AAW24063 standard; Protein; 241 AA.
XX AC AAW24063;
XX DT 17-MAR-1998 (first entry)
XX DE Human WSX receptor agonist antibody clone #17.
XX KW Human; WSX receptor; clone #17; identification; purification;
XX KW ligand; activator; antibody; agonist; proliferation; obesity;
XX KW differentiation; anaemia; treatment; neoplasia; arteriosclerosis;
XX KW Type II diabetes; polycystic ovarian disease;
XX KW cardiovascular disease; osteoarthritis; dermatological disorder;
XX KW hypertension; insulin resistance; hypercholesterolaemia;
XX KW hypertriglyceridaemia; cancer; cholelithiasis.
XX OS Homo sapiens.
XX PN WO9725425-A1.
XX PD 17-JUL-1997.
XX PF 07-JAN-1997; 97WO-US00325.
XX PR 20-JUN-1996; 96US-0667197.
XX PR 08-JAN-1996; 96US-0585005.
XX PA (GETH ) GENENTECH INC.
XX PI Bennett B, Carter PJ, Chiang NY, Kim KJ, Matthews W;
XX PI Rodrigues ML;
XX DR WPI; 1997-372864/34.
XX WSX*receptor and related antibodies and ligands - used to develop
XX PT products for diagnosis and therapy, e.g. for improving
XX PT haematopoiesis or for treating tumours

```

---

```

XX PS Example 14; Pages 122-123; 219pp; English.
XX CC The present sequence is an agonist antibody clone to the human WSX
XX CC receptor, which can be used to identify and purify ligands and
XX CC activators. An anti-WSX receptor antibody can be used as an agonist
XX CC to activate the WSX receptor, leading to enhanced proliferation or
XX CC differentiation of a cell expressing the WSX receptor. It can also
XX CC be used to decrease body weight and/or fat-depot weight and/or food
XX CC intake in an obese mammal. WSX receptor ligands can be used to
XX CC enhance proliferation or differentiation of lymphoid, myeloid or
XX CC erythroid blood cell lineages. This is useful when a mammal,
XX CC especially a human, is suffering from decreased blood cell levels,
XX CC i.e. anaemia, caused by chemotherapy, radiation therapy or bone
XX CC marrow transplantation therapy. It can also be used to repopulate
XX CC blood cells in a mammal. The products can also be used to treat,
XX CC e.g. neoplastic disorders, arteriosclerosis, type II diabetes,
XX CC polycystic ovarian disease, cardiovascular diseases,
XX CC osteoarthritis, dermatological disorders, hypertension, insulin
XX CC resistance, hypercholesterolaemia, hypertriglyceridaemia, cancer
XX CC and cholelithiasis.
XX SQ Sequence 241 AA;
    Query Match 75.0%; Score 36; DB 18; Length 241;
    Best Local Similarity 66.7%; Pred. No. 30;
    Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 QQLVVEYPPT 9
Db 223 QQLISYPLT 231
    | | | | |
    | | | | |

Search completed: April 22, 2003, 12:51:13
Job time : 23.8214 secs

```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:53:59 ; Search time 9.48214 Seconds  
(without alignments)  
76.055 Million cell updates/sec

Title: US-09-674-716B-7

Perfect score: 48

Sequence: 1 QOLVEYPFT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/prodata/1/pubpaa/US08\_NEW PUB.pdb:\*
  - 2: /cgn2\_6/prodata/1/pubpaa/US06\_NEW PUB.pdb:\*
  - 3: /cgn2\_6/prodata/1/pubpaa/US05\_NEW PUB.pdb:\*
  - 4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pdb:\*
  - 5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW PUB.pdb:\*
  - 6: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pdb:\*
  - 7: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pdb:\*
  - 8: /cgn2\_6/prodata/1/pubpaa/US09\_NEW PUB.pdb:\*
  - 9: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pdb:\*
  - 10: /cgn2\_6/prodata/1/pubpaa/US10\_NEW PUB.pdb:\*
  - 11: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pdb:\*
  - 12: /cgn2\_6/prodata/1/pubpaa/US60\_NEW PUB.pdb:\*
  - 13: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pdb:\*
  - 14: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pdb:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	77.1	100	10	US-09-840-459-32
2	36	75.0	241	1	US-08-779-457-50
3	36	75.0	251	9	US-09-880-748-1310
4	36	75.0	253	9	US-09-880-748-936
5	33	68.8	82	10	US-09-864-761-47635
6	33	68.8	107	9	US-09-848-798-40
7	33	68.8	726	10	US-09-973-451-10
8	32	66.7	105	9	US-10-164-359-4
9	32	66.7	105	10	US-09-107-058-8
10	32	66.7	105	10	US-09-761-117-8
11	32	66.7	109	9	US-09-144-886-97
12	32	66.7	295	9	US-10-045-545-1
13	32	66.7	296	9	US-10-045-545-1
14	32	66.7	296	9	US-10-045-545-3
15	31	64.6	9	10	US-09-976-787-6
16	31	64.6	9	10	US-09-865-198-6
17	31	64.6	35	9	US-09-809-391-745
18	31	64.6	106	10	US-09-976-787-24
19	31	64.6	106	10	US-09-865-198-23

20	64.6	107	9	US-09-144-886-75	Sequence 75, Appli
21	64.6	108	10	US-09-976-787-8	Sequence 8, Appli
22	64.6	108	10	US-09-865-198-8	Sequence 8, Appli
23	64.6	112	10	US-09-764-877-2008	Sequence 2008, Ap
24	64.6	131	10	US-09-881-823-6	Sequence 6, Appli
25	64.6	169	10	US-09-815-242-5756	Sequence 5756, Ap
26	64.6	172	10	US-09-815-242-12613	Sequence 12613, A
27	64.6	172	10	US-09-815-242-12852	Sequence 12852, A
28	64.6	238	10	US-09-976-787-29	Sequence 28, Appli
29	64.6	238	10	US-09-865-198-28	Sequence 28, Appli
30	64.6	240	10	US-09-976-787-28	Sequence 28, Appli
31	64.6	240	10	US-09-865-198-27	Sequence 27, Appli
32	64.6	366	10	US-09-901-252-17	Sequence 17, Appli
33	64.6	380	10	US-09-098-079-25	Sequence 25, Appli
34	64.6	413	9	US-09-809-391-751	Sequence 751, App
35	64.6	734	9	US-09-738-626-5405	Sequence 5405, Ap
36	64.6	806	9	US-10-025-380-1117	Sequence 1117, Ap
37	64.6	806	9	US-10-025-380-1118	Sequence 1118, Ap
38	64.6	806	10	US-09-922-217-1117	Sequence 1117, Ap
39	64.6	806	10	US-09-922-217-1118	Sequence 1118, Ap
40	62.5	62	9	US-09-828-708-51	Sequence 51, Appl
41	62.5	62	9	US-09-309-196-116	Sequence 116, App
42	62.5	104	10	US-09-828-708-2	Sequence 2, Appli
43	62.5	168	10	US-09-764-864-1223	Sequence 1223, Ap
44	62.5	238	10	US-09-815-242-5463	Sequence 5463, Ap
45	62.5	242	10	US-09-815-242-12164	Sequence 12164, A

## ALIGNMENTS

RESULT 1

US-09-840-459-32

Sequence 32, Application US/09840459

Patent No. US20020150576A1

GENERAL INFORMATION:

APPLICANT: Lakosa, Gregory J.

APPLICANT: Horvath, Christopher

APPLICANT: Newman, Walter

APPLICANT: Jones, S. Tarran

APPLICANT: O'Brien, Siobhan H.

APPLICANT: O'Keefe, Theresa

TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND METHODS OF USE THEREFOR

FILE REFERENCE: 1855.1052-012

CURRENT APPLICATION NUMBER: US/09/840,459

CURRENT FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: PCT/US01/03537

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 09/497,625

PRIOR FILING DATE: 2000-02-03

PRIOR APPLICATION NUMBER: 09/359,193

PRIOR FILING DATE: 1999-07-22

PRIOR APPLICATION NUMBER: 09/121,781

PRIOR FILING DATE: 1998-07-23

NUMBER OF SEQ ID NOS: 107

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 32

LENGTH: 100

TYPE: PRT

ORGANISM: Mus musculus

US-09-840-459-32

Query Match 77.1%; Score 37; DB 10; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QOLVEYP 7

Db 94 QOLVEYP 100

RESULT 2

US-08-779-457-50  
; Sequence 50, Application US/08779457  
; Publication No. US2002019357A1  
; GENERAL INFORMATION:  
; APPLICANT: Carter, Paul J.  
; APPLICANT: Chang, Nancy Y.  
; APPLICANT: Kyung, Jin Kim  
; APPLICANT: Matthews, William  
; APPLICANT: Rodriguez, Maria L.  
; TITLE OF INVENTION: WSK RECEPTOR AGONIST ANTIBODIES  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/779,457  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/667197  
; FILING DATE: 06/20/96  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/585005  
; FILING DATE: 01/08/96  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0986P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 241 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-779-457-50

Query Match 75.0%; Score 36; DB 1; Length 241;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9  
|||: |||  
Db 223 QQLISYPLT 231

## RESULT 3

US-09-880-748-1310  
; Sequence 1310, Application US/09880748  
; Publication No. US2003005937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1310  
; LENGTH: 251  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1310

Query Match 75.0%; Score 36; DB 9; Length 251;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9  
|||: |||  
Db 232 QQLISYPLT 240

## RESULT 4

US-09-880-748-936  
; Sequence 936, Application US/09880748  
; Publication No. US2003005937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 936  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-936

Query Match 75.0%; Score 36; DB 9; Length 253;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9  
|||: |||  
Db 234 QQLISYPLT 242

## RESULT 5

US-09-864-761-47635  
; Sequence 47635, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Acomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312

```
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47635
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL009031.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.1
; OTHER INFORMATION: EST HUMAN HIT: AW612517.1, EVALUE 7.00e-43
; OTHER INFORMATION: SWISSPROT HIT: P54253, EVALUE 1.00e-43
US-09-864-761-47635

Query Match      68.8%; Score 33; DB 10; Length 82;
Best Local Similarity 100.0%; Pred. No. 18;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      3 LVEYFF 8
      |||||
Db      6 LVEYFF 11

RESULT 6
US-09-848-798-40
; Sequence 40, Application US/09848798
; Publication No. US20030040605A1
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: RA(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/848,798
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/240,274
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/028,550
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain I09
US-09-848-798-40

Query Match      68.8%; Score 33; DB 9; Length 107;
Best Local Similarity 66.7%; Pred. No. 23;
Matches      6; Conservative      1; Mismatches      2; Indels      0; Gaps      0;

Qy      1 QQLVEYFP 9
      |||||
Db      88 QQLNSYPT 96

RESULT 7
US-09-973-451-10
; Sequence 10, Application US/09973451
; Patent No. US20020132328A1
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY (ADP-RIBOSE) GLYCOHYDROLASE
; TITLE OF INVENTION: (PARG) ENZYMES,
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/973,451
; CURRENT FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US/09/302,812
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 60/083,768
; PRIOR FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 10
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-973-451-10

Query Match      68.8%; Score 33; DB 10; Length 726;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches      5; Conservative      2; Mismatches      0; Indels      0; Gaps      0;

Qy      1 QQLVEYFP 7
      |||||
Db      172 QQLIEYP 178

RESULT 8
US-10-164-359-4
; Sequence 4, Application US/10164359
; Publication No. US20030012776A1
; GENERAL INFORMATION:
; APPLICANT: Chin, Khew-Voon
; TITLE OF INVENTION: Nucleic Acid and Protein Expressed Thereby and Their Involvement
; FILE OF INVENTION: Stress
; FILE REFERENCE: 601-1-108US
; CURRENT APPLICATION NUMBER: US/10/164,359
; CURRENT FILING DATE: 2002-08-06
; PRIOR APPLICATION NUMBER: PCT/US00/33438
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 60/169,418
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
```

```
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-359-4

Query Match          66.7%; Score 32; DB 9; Length 105;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9
Db 66 QQILEYAYT 74

RESULT 9
US-09-107-058-8
; Sequence 8, Application US/09107058
; Patent No. US20010010922A1
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; APPLICANT: Niu, Hui-Feng
; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC
; LOCUS bcl-6
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/107,058
; CLASSIFICATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-761-117-8

Query Match          66.7%; Score 32; DB 10; Length 105;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9
Db 72 QQILEYAYT 80

RESULT 11
US-09-144-886-97
; Sequence 97, Application US/09144886
; Patent No. US20020155114A1
; GENERAL INFORMATION:
; APPLICANT: Marks, James D
; APPLICANT: Amerdorfer, Peter
; TITLE OF INVENTION: Therapeutic Monoclonal Antibodies That Neutralize
; Botulinum Neurotoxins
; FILE REFERENCE: 2500.117USO
; CURRENT APPLICATION NUMBER: US/09/144,886
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 97
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: BoNT/A clone
; OTHER INFORMATION: 1F3 region VL epitope 4
US-09-144-886-97

Query Match          66.7%; Score 32; DB 9; Length 109;
Best Local Similarity 55.6%; Pred. No. 36;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 QQLVEYPFT 9
DB 90 QQYIGYPT 98

RESULT 12
US-10-045-545-4
; Sequence 4, Application US/10045545
; Publication No. US20030027124A1
; GENERAL INFORMATION:
; APPLICANT: Maines, Mahin D.
; TITLE OF INVENTION: METHODS OF MODIFYING CELL STRUCTURE AND REMODELING
; FILE REFERENCE: 176/60981
; CURRENT APPLICATION NUMBER: US/10/045,545
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: 60/261,500
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-045-545-4

Query Match 66.7%; Score 32; DB 9; Length 295;
Best Local Similarity 85.7%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVEYPFT 9
DB 94 LVEYPMT 100

RESULT 13
US-10-045-545-1
; Sequence 1, Application US/10045545
; Publication No. US20030027124A1
; GENERAL INFORMATION:
; APPLICANT: Maines, Mahin D.
; TITLE OF INVENTION: METHODS OF MODIFYING CELL STRUCTURE AND REMODELING
; FILE REFERENCE: 176/60981
; CURRENT APPLICATION NUMBER: US/10/045,545
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: 60/261,500
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-045-545-1

Query Match 66.7%; Score 32; DB 9; Length 296;
Best Local Similarity 85.7%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVEYPFT 9
DB 95 LVEYPMT 101

RESULT 14
US-10-045-545-3
; Sequence 3, Application US/10045545
; Publication No. US20030027124A1
; GENERAL INFORMATION:
; APPLICANT: Maines, Mahin D.
; TITLE OF INVENTION: METHODS OF MODIFYING CELL STRUCTURE AND REMODELING
; FILE REFERENCE: 176/60981
; CURRENT APPLICATION NUMBER: US/10/045,545
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: 60/261,500
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-045-545-3

Query Match 66.7%; Score 32; DB 9; Length 296;
Best Local Similarity 85.7%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVEYPFT 9
DB 95 LVEYPMT 101

RESULT 15
US-09-976-787-6
; Sequence 6, Application US/09976787
; Patent No. US20020064528A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Zhenping
; APPLICANT: Witte, Larry
; TITLE OF INVENTION: Antibodies Specific to KDR and Uses Thereof
; FILE REFERENCE: 11245/46505
; CURRENT APPLICATION NUMBER: US/09/976,787
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 09/493,539
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,726
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 6
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-976-787-6

Query Match 64.6%; Score 31; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.7e+05;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9
DB 1 QQRSSYPFT 9

Search completed: April 22, 2003, 13:11:13
Job time : 10.4821 secs
```





GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:24 ; Search time 4.91071 Seconds  
(without alignments)  
97.882 Million cell updates/sec

Title: US-09-674-716B-9

Perfect score: 33

Sequence: 1 GYWS 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.\*

2: Pir2.\*

3: Pir3.\*

4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	67	2 A84376	hypothetical prote
2	33	100.0	117	1 G2MSU1	Ig heavy chain V r
3	33	100.0	444	2 S09681	citrate transport
4	33	100.0	1043	2 F97302	hypothetical prote
5	30	90.9	86	2 T42185	conserved hypothet
6	30	90.9	298	2 AE1752	hypothetical prote
7	30	90.9	298	2 AH1728	Abid phage protein
8	30	90.9	315	2 H86836	carbamate kinase (
9	30	90.9	348	2 G69142	GDP-D-mannose dehy
10	30	90.9	358	2 H83554	hypothetical prote
11	30	90.9	375	2 E83163	hypothetical prote
12	30	90.9	417	2 T08724	hypothetical prote
13	30	90.9	448	2 E95249	PTS system, IIC co
14	30	90.9	448	2 G98113	conserved hypothet
15	30	90.9	452	2 T00113	undecaprenyl-phosp
16	30	90.9	464	2 T28818	hypothetical prote
17	30	90.9	475	2 T46745	arginine/ornithine
18	30	90.9	476	2 D84505	probable membrane
19	30	90.9	480	2 S39978	scra protein - sta
20	30	90.9	480	2 D90038	PTS system, sucros
21	30	90.9	490	2 C86879	arginine/ornithine
22	30	90.9	518	1 G69804	multidrug-efflux t
23	30	90.9	535	2 S78598	D-ribulokinase (BC
24	30	90.9	583	2 C72544	probable glycol-tr
25	30	90.9	592	1 IKEBCA	colicin A - Citrob
26	30	90.9	592	2 T15600	hypothetical prote
27	30	90.9	608	2 JQ1462	phosphoenolpyruvat
28	30	90.9	667	1 VCLJGL	env polyprotein pr
29	30	90.9	847	2 JC4836	alpha-glucuronidas

#### ALIGNMENTS

##### RESULT 1

A84376

hypothetical protein Vng2253h [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: A84376

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li.

A:Title: Genome sequence of Halobacterium species NRC-1

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: A84376

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-67 <STO>

A:Cross-references: GB:AE004437; NID:g10581666; PIDN:AGG20373.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG2253H

Query Match 100.0%; Score 33; DB 2; Length 67;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5

Db 59 GYWS 63

##### RESULT 2

G2MSU1

Ig heavy chain V region (UPC10) - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-Aug-1984 #sequence\_revision 03-Aug-1984 #text\_change 23-Mar-2001

C:Accession: A02081

R:Auffray, C.; Slikov, J.L.; Ollo, R.; Rougeon, P.

Ann. Immunol. (Paris) 132D, 77-88, 1981

A:Title: Correlation between D region structure and antigen-binding specificity: evidence:

A:Reference number: A02081; MUID:83021113; PMID:6181731

A:Accession: A02081

A:Molecule type: mRNA

A:Residues: 1-117 <AUF>

C:Comment: This chain was isolated from an Ig gamma-2a myeloma protein binding 2,6-leval;

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotrimer; immunoglobulin

F;15-98/Domain: immunoglobulin homology <IMM>

F;22-96/Disulfide bonds: #status predicted

Query Match 100.0%; Score 33; DB 1; Length 117;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5  
 |||||  
 Db 31 GYWS 35

# RESULT 3

S09681  
 citrate transport protein - Klebsiella pneumoniae plasmid pBS1

N;Alternate names: citrate carrier  
 C;Species: Klebsiella pneumoniae  
 C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 26-Aug-1999  
 C;Accession: S09681  
 R;van der Rest, M.E.; Schwarz, E.; Oesterhelt, D.; Konings, W.N.  
 Eur. J. Biochem. 189, 401-407, 1990  
 A;Title: DNA sequence of a citrate carrier of Klebsiella pneumoniae.  
 A;Reference number: S09681; MUID:90249385; PMID:2186908  
 A;Accession: S09681  
 A;Molecule type: DNA  
 A;Residues: 1-444 <RES>  
 A;Cross-references: EMBL:X51479; NID:943800; PIDN:CAA35844.1; PID:943801  
 C;Genetics:  
 A;Genome: plasmid pBS1  
 C;Superfamily: citrate utilization determinant  
 C;Keywords: membrane protein

Query Match 100.0%; Score 33; DB 2; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5  
 |||||  
 Db 415 GYWS 419

# RESULT 4

P97302  
 hypothetical protein CAC3275 [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum  
 C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C;Accession: F97302  
 R;Nolling, J.; Bretton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
 A;Reference number: A96900; MUID:21359325; PMID:21359325  
 A;Accession: F97302  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1043 <KUR>  
 A;Cross-references: GB:AE001437; PIDN:AAK81209.1; PID:gl5026351; GSPDB:GN00168  
 A;Experimental source: Clostridium acetobutylicum ATCC824  
 C;Genetics:  
 A;Gene: CAC3275

Query Match 100.0%; Score 33; DB 2; Length 1043;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5  
 |||||  
 Db 580 GYWS 584

# RESULT 5

T42185  
 conserved hypothetical protein L7085 - Escherichia coli plasmid pO157

C;Species: Escherichia coli  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C;Accession: T42185  
 R;Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.  
 Nucleic Acids Res. 26, 4196-4204, 1998  
 A;Title: The complete DNA sequence and analysis of the large virulence plasmid of Esche-

A;Reference number: Z22068; MUID:98391744; PMID:9722640  
 A;Accession: T42185  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-86 <EUR>  
 A;Cross-references: EMBL:AF074613; PIDN:AA070153.1  
 A;Experimental source: strain EDJ933; serotype O157:H7  
 C;Genetics:  
 A;Genome: plasmid pO157  
 A;Note: L7085

Query Match 90.9%; Score 30; DB 2; Length 86;  
 Best Local Similarity 80.0%; Pred. No. 69;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5  
 |||||  
 Db 72 GYWS 76

# RESULT 6

AE1752  
 hypothetical protein homolog lin2562 [imported] - Listeria innocua (strain Clip11262)  
 C;Species: Listeria innocua  
 C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C;Accession: AE1752  
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Faihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
 A;Reference number: AE1077; MUID:21537279; PMID:11679669  
 A;Accession: AE1752  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-298 <GLA>  
 A;Cross-references: GB:AL592022; PIDN:CAC97789.1; PID:gl6415084; GSPDB:GN00178  
 A;Experimental source: strain Clip11262  
 C;Genetics:  
 A;Gene: lin2562

Query Match 90.9%; Score 30; DB 2; Length 298;  
 Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5  
 |||||  
 Db 45 GYWS 49

# RESULT 7

AH1728  
 AbD phage protein homolog lin2373 [imported] - Listeria innocua (strain Clip11262)  
 C;Species: Listeria innocua  
 C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C;Accession: AH1728  
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Faihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AH1728  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-298 <GLA>  
 A;Cross-references: GB:AL592022; PIDN:CAC97600.1; PID:gl6414896; GSPDB:GN00178  
 A;Experimental source: strain Clip11262  
 C;Genetics:

A:Gene: lin2373

Query Match 90.9%; Score 30; DB 2; Length 298;  
Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:  
Db 45 GYWLS 49

#### RESULT 8

H86836  
C:Species: Lactococcus lactis subsp. lactis (strain IL14)  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: H86836  
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: H86836  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-315 <STO>  
A:Cross-references: GB:AB005176; PID:g12724712; PIDN:AAK05794.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: arcC3  
C:Superfamily: carbamate kinase  
C:Keywords: phosphotransferase

Query Match 90.9%; Score 30; DB 2; Length 315;  
Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:  
Db 88 GYWMA 92

#### RESULT 9

G69142  
GDP-D-mannose dehydratase - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
C:Accession: G69142  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qi, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: G69142  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-348 <MTH>  
A:Cross-references: GB:AE000818; GB:AE000666; NID:g2621384; PIDN:AAB84839.1; PID:g262138  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH333  
C:Superfamily: GDP-D-mannose dehydratase

Query Match 90.9%; Score 30; DB 2; Length 348;  
Best Local Similarity 80.0%; Pred. No. 2.6e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:  
Db 156 GYWT 160

#### RESULT 10

H83554

hypothetical protein PA0736 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: H83554  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: H83554  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <STO>  
A:Cross-references: GB:AE004508; GB:AE004091; NID:g9946611; PIDN:AAG04125.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA0736

Query Match 90.9%; Score 30; DB 2; Length 358;  
Best Local Similarity 80.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:  
Db 203 GYWLS 207

#### RESULT 11

E83163  
hypothetical protein PA3863 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: E83163  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: E83163  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-375 <STO>  
A:Cross-references: GB:AE004803; GB:AE004091; NID:g9950035; PIDN:AAG07250.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3863

Query Match 90.9%; Score 30; DB 2; Length 375;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:  
Db 19 GYWLS 23

#### RESULT 12

T08724  
hypothetical protein DKFP566D213.1 - human  
C:Species: Homo sapiens (man)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 13-Aug-1999  
C:Accession: T08724  
R:Koeher, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, May 1999  
A:Reference number: Z16468  
A:Accession: T08724  
A:Molecule type: mRNA  
A:Residues: 1-417 <KOE>  
A:Cross-references: EMBL:AL050275  
A:Experimental source: fetal kidney; clone DKFP566D213

C:Genetics:  
A:Note: DKP2p566D213.1

Query Match 90.9%; Score 30; DB 2; Length 417;  
Best Local Similarity 80.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:|  
Db 402 GYWLS 406

RESULT 13  
B95249

PTS system, IIC component, probable [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001

C:Accession: B95249

R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,  
nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: B95249

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-448 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK76187.1; PID:g14973641; GSPDB:GN00164; TIGR:SP4

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP2129

Query Match 90.9%; Score 30; DB 2; Length 448;  
Best Local Similarity 80.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:|  
Db 186 GYWLS 190

RESULT 14  
G98113

conserved hypothetical protein spr1938 [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001

C:Accession: G98113

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, F.; McAhren, S.; M  
Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: G98113

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-448 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAAL00740.1; PID:g15459636; GSPDB:GN00174

C:Genetics:

A:Gene: spr1938

Query Match 90.9%; Score 30; DB 2; Length 448;  
Best Local Similarity 80.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:|  
Db 186 GYWLS 190

RESULT 15

T00113

undecaprenyl-phosphate galactosephosphotransferase homolog - Actinobacillus actinomycete

C:Species: Actinobacillus actinomycetemcomitans

C>Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 21-Jul-2000

C:Accession: T00113

R:Nakano, Y.; Yoshida, Y.; Yamashita, Y.; Koga, T.

Biochim. Biophys. Acta 1442, 409-414, 1998

A:Title: A gene cluster for 6-deoxy-L-talan synthesis in Actinobacillus actinomycetemcom

A:Reference number: Z14111; MUID:99023768; PMID:9805002

A:Accession: T00113

A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-452 <NAK>

A:Cross-references: EMBL:AB010415; NID:G3132248; PIDN:BAA28142.1; PID:G3132265

A:Experimental source: strain NCTC 9710

C:Superfamily: xps2A protein

Query Match 90.9%; Score 30; DB 2; Length 452;  
Best Local Similarity 80.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:|  
Db 11 GYWLS 15

Search completed: April 22, 2003, 12:54:51  
Job time : 6.91071 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:46:53 ; Search time 2.5 seconds  
(without alignments)  
82.953 Million cell updates/sec

Title: US-09-674-716B-9

Perfect score: 33

Sequence: 1 GYWS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	33	100.0	117	1 HV41_MOUSE	P01811 mus musculu
2	33	100.0	444	1 CITI_KLEPN	P16482 klebsiella
3	30	90.9	214	1 CAPA_KLULA	O74232 kluyveromy
4	30	90.9	292	1 AMPR_PROST	O69772 providencia
5	30	90.9	309	1 ARCC_ENTFC	P35836 enterococcu
6	30	90.9	475	1 ARCC_LACSK	O53092 lactobacill
7	30	90.9	478	1 ARCD_CLOPE	O46170 clostridium
8	30	90.9	480	1 PTSE_STAXY	P51184 staphylococ
9	30	90.9	583	1 SYG_AERPE	O975f8 aeropyrum p
10	30	90.9	592	1 CEA_CITFR	P04480 citrobacter
11	30	90.9	608	1 PPCK_NEOFR	P22130 neocallimas
12	30	90.9	667	1 ENV_GALV	P21415 gibbon ape
13	30	90.9	847	1 AGUA_TIRE	O99024 trichoderma
14	30	90.9	936	1 CAPP_RHOMR	O59757 rhodothermu
15	30	90.9	1063	1 CZCA_ALCEU	P13511 alcaligenes
16	30	90.9	1063	1 CZCA_ALCSP	P94177 alcaligenes
17	30	90.9	1502	1 MRPE_RAT	O88269 rattus norv
18	30	90.9	1503	1 MRPE_HUMAN	O95255 homo sapien
19	29	87.9	252	1 PRIO_ATEPA	P51446 ateles pani
20	29	87.9	252	1 PRIO_RABIT	O95211 oryctolagus
21	29	87.9	270	1 GPR1_YARLI	P41943 yarrowia li
22	29	87.9	271	1 YAD8_SCHPO	O09834 schizosacch
23	29	87.9	287	1 LEPA_LEGPN	O68433 legionella
24	29	87.9	314	1 ARCC_LACSK	O53090 lactobacill
25	29	87.9	388	1 Y535_BUCAL	P57601 buchnera ap
26	29	87.9	395	1 ARP2_DROME	P45888 drosophila
27	29	87.9	482	1 ARCD_PSEAE	P18275 pseudomonas
28	29	87.9	494	1 AMY1_SACFI	P21567 saccharomyc
29	29	87.9	503	1 PR19_YEAST	P32523 saccharomyc
30	29	87.9	512	1 AMY1_DEBOC	P19269 debaryomyce
31	29	87.9	518	1 RTN_ECOLI	P76446 escherichia
32	29	87.9	533	1 GLPT_HUMAN	P57057 homo sapien
33	29	87.9	579	1 YGT6_YEAST	P53099 saccharomyc

34	29	87.9	587	1 PH84_YEAST	P25297 saccharomyc
35	29	87.9	598	1 CSTA_BACSU	P94532 bacillus su
36	29	87.9	615	1 SPPA_HAEIN	P45243 haemophilus
37	29	87.9	618	1 SPPA_ECOLI	P08395 escherichia
38	29	87.9	639	1 Y119_MYCTU	Q50614 mycobacteri
39	29	87.9	680	1 GAOA_DACDE	Q01745 daetylum d
40	29	87.9	774	1 YQV5_CAEEL	Q09554 caenorhabdi
41	29	87.9	776	1 VP4_ROT99	Q06894 canine rota
42	29	87.9	776	1 VP4_ROT96	Q07416 feline rota
43	29	87.9	776	1 VP4_ROT9Y	Q06895 human rotav
44	28	84.8	141	1 YEDD_SALTY	Q06399 salmonella
45	28	84.8	184	1 MUP3_MOUSE	P04939 mus musculu

#### ALIGNMENTS

RESULT 1					
HV41_MOUSE					
ID	HV41_MOUSE	STANDARD;	PRT;	117 AA.	
AC	P01811;				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Ig heavy chain V region UPC10.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=83021113; PubMed=6181731;				
RA	Auffray C., Sikorav J.L., Ollo R., Rougeon F.;				
RT	"Correlation between D region structure and antigen-binding				
RT	specificity: evidences from the comparison of closely related				
RT	immunoglobulin VH sequences."				
RL	Ann. Immunol. (Paris) 132D:77-88(1981).				
CC	-1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IG GAMMA-2A MYELOMA				
CC	PROTEIN BINDING 2,6-LEVAN.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch)				
CC	-----				
DR	EMBL; M36631; AAA38078.1; -				
DR	FIR; A02081; G2MSU1.				
DR	HSSP; P01810; 2FBJ.				
DR	InterPro; IPR003006; Ig_MHC.				
DR	InterPro; IPR003596; Ig_v.				
DR	Pfam; PF00047; Ig_1.				
DR	SMART; SM00406; IGV; 1.				
KW	Immunoglobulin V region.				
FT	NON_TER	117	117		
SQ	SEQUENCE	117 AA;	13001 MW;	B20A1074F8E99E7F CRC64;	
Query Match 100.0%; Score 33; DB 1; Length 117;					
Best Local Similarity 100.0%; Pred. No. 9.5;					
Matches 5; Conservative 0; Mismatches 0; Gaps 0;					
Oy 1 GYWS 5					
Db 31 GYWS 35					
-----					
RESULT 2					
CITI_KLEPN					
ID	CITI_KLEPN	STANDARD;	PRT;	444 AA.	
AC	P16482;				
DT	01-AUG-1990	(Rel. 15, Created)			

DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Citrate-proton symporter (Citrate transporter) (Citrate carrier  
 DE protein)  
 GN CITH OR CIT.  
 OS Klebsiella pneumoniae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Klebsiella.  
 RN NCBI\_TaxID=573;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13882;  
 RX MEDLINE=90249385; PubMed=2186908;  
 RA van der Rest M.E., Schwarz E., Oosterhelt D., Konings W.N.;  
 RT "DNA sequence of a citrate carrier of Klebsiella pneumoniae";  
 RL Eur. J. Biochem. 189:401-407(1990).  
 CC -!- FUNCTION: UPTAKE OF CITRATE ACROSS THE BOUNDARY MEMBRANE WITH  
 CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).  
 CC -!- PATHWAY: ALLOWS THE UTILIZATION OF CITRATE AS A SOLE SOURCE OF  
 CC CARBON AND ENERGY.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -!- SIMILARITY: TO CIT IN TN3411 AND PWR60 FOUND IN E.COLI AND TO  
 CC -!- CIT IN CITROBACTER AMALONATICUS.  
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X51479; CAA35944.1; -;  
 DR PIR; S09681; S09681.  
 DR InterPro; IPR004736; Cit H symport.  
 DR InterPro; IPR003662; Sub transporter.  
 DR Pfam; PF00083; sugar tr; 1.  
 DR TIGRfam; TIGR00883; 2A0106; 1.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 2.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 DR Transprot; Transmembrane; Inner membrane; Citrate utilization;  
 KW Symport.  
 FT DOMAIN 1 41  
 FT TRANSMEM 42 62  
 FT DOMAIN 63 72  
 FT TRANSMEM 73 93  
 FT DOMAIN 94 114  
 FT TRANSMEM 115 135  
 FT DOMAIN 136 137  
 FT TRANSMEM 138 158  
 FT DOMAIN 159 177  
 FT TRANSMEM 178 198  
 FT DOMAIN 199 199  
 FT TRANSMEM 200 220  
 FT DOMAIN 221 251  
 FT TRANSMEM 252 272  
 FT DOMAIN 273 289  
 FT TRANSMEM 290 310  
 FT DOMAIN 311 318  
 FT TRANSMEM 319 339  
 FT DOMAIN 340 340  
 FT TRANSMEM 341 361  
 FT DOMAIN 362 379  
 FT TRANSMEM 380 400  
 FT DOMAIN 401 411  
 FT TRANSMEM 412 432  
 FT DOMAIN 433 444  
 SQ SEQUENCE 444 AA; 48142 MW; D431F732EED3EC2 CRC64;  
 Query Match 100.0%; Score 33; DB 1; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5  
 DB 415 GYWS 419  
 RESULT 3  
 ID CAPA\_KLULA STANDARD; PRT; 214 AA.  
 AC O74232;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE F-actin capping protein alpha subunit (Fragment).  
 GN CAP1.  
 OS Kluyveromyces lactis (Yeast).  
 OC Eukaryota; Fungi; Ascomycota;  
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
 OX NCBI\_TaxID=28985;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Banfield D.K.;  
 RT "DNA sequence of the SFT1 gene from Kluyveromyces lactis";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DDBB databases.  
 CC -!- FUNCTION: F-ACTIN CAPPING PROTEINS BIND IN A CA(2+)-INDEPENDENT  
 CC MANNER TO THE FAST GROWING ENDS OF ACTIN FILAMENTS (BARBED END)  
 CC THEREBY BLOCKING THE EXCHANGE OF SUBUNITS AT THESE ENDS. UNLIKE  
 CC OTHER CAPPING PROTEINS (SUCH AS GELSOLIN AND SEVERIN), THESE  
 CC PROTEINS DO NOT SEVER ACTIN FILAMENTS (BY SIMILARITY).  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE F-ACTIN CAPPING PROTEIN ALPHA SUBUNIT  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF072674; AAC24885.1; -;  
 DR InterPro; IPR002189; F-actin cap A.  
 DR Pfam; PF01267; F-actin cap A; 1.  
 DR PRINTS; PR00191; F-actinCAPA.  
 DR ProDom; PD006960; F-actin cap A; 1.  
 DR PROSITE; PS00748; F-ACTIN\_CAPPING\_A\_1; 1.  
 DR PROSITE; PS00749; F-ACTIN\_CAPPING\_A\_2; 1.  
 KW Actin-binding; Capping protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 214 AA; 24380 MW; 31851D6573630C02 CRC64;  
 Query Match 90.9%; Score 30; DB 1; Length 214;  
 Best Local Similarity 80.0%; Pred. No. 60;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWS 5  
 DB 111 GYWS 115  
 RESULT 4  
 ID AMPR\_PROST STANDARD; PRT; 292 AA.  
 AC O69772;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Transcriptional activator protein ampr.  
 GN AMPR.  
 OS Providencia stuartii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```

RL  J. Mol. Biol. 235:1345-1347(1994)
CC  -| CATALYTIC ACTIVITY: ATP + NH(3) + CO(2) = ADP + carbamoyl
CC  phosphate.
CC  -| PATHWAY: Arginine degradation via arginine deiminase; third step.
CC  -| SUBUNIT: HOMODIMER (PREDOMINANTLY) AND HOMOTETRAMER.
CC  -| SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC  -| SIMILARITY: BELONGS TO THE CARBAMATE KINASE FAMILY.
CC  -----
CC  * This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as the content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AJ2233331; CRA11270.1; -
DR  EMBL; AJ2233332; CRA11271.1; -
DR  PIR; S44002; S44002.
DR  HSP; P95474; 1E19.
DR  InterPro; IPR001048; Aa_kinase.
DR  InterPro; IPR003964; Bac_carb_kinase.
DR  Pfam; PF00696; aakinae; 1.
DR  TIGRfam; TIGR00746; arcC; 1.
KW  Transferase; Kinase; Arginine metabolism.
FT  INIT_MET 0
FT  CONFLICT 33 157 MISSING (IN REF. 2).
FT  CONFLICT 168 168 I -> D (IN REF. 2).
SQ  SEQUENCE 309 AA; 32795 MW; 5014514A7FEEB290 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 309;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWMS 5
DB 86 GYWLS 90

```

```

RESULT 6
RCRD_LACSK
ID _ARCD_LACSK STANDARD; PRF; 475 AA.
AC O53092;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Arginine/ornithine antiporter.
GN ARCD.
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OC NCBI_TaxID=1599;
[1]
SEQUENCE FROM N.A.
MEDLINE=98361904; PubMed=3696763;
RX Zuniga M., Champomier-Verges M., Zagorec M., Perez-Martinez G.;
RA "Structural and functional analysis of the gene cluster encoding the
RT enzymes of the arginine deiminase pathway of Lactobacillus sake.";
RL J. Bacteriol. 180:4154-4159(1998).
CC -1- FUNCTION: CATALYZES AN ELECTRONEUTRAL EXCHANGE BETWEEN ARGinine
AND ORNITHINE TO ALLOW HIGH-EFFICIENCY ENERGY CONVERSION IN THE
CC ARGinine DEIMINASE PATHWAY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
CC PERMEASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).

```

```
CC EMBL; AJ001330; CAA04686.1; -.
DR InterPro; IPR002293; AA/rel_prmease1.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00305; 2A0302; 1.
KW Transport; Antiport; Amino-acid transport; Transmembrane.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 74 94 POTENTIAL.
FT TRANSMEM 101 121 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 205 225 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 283 303 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 397 417 POTENTIAL.
FT TRANSMEM 451 471 POTENTIAL.
SQ SEQUENCE 475 AA; 51881 MW; 8591A01F6A2203CC CRC64;

Query Match 90.9%; Score 30; DB 1; Length 475;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWSMS 5
DB 94 GYWSLS 98

RESULT 7
ID ARCD CLOPE STANDARD; PRT; 478 AA.
AC Q46170.
DT 15-DEC-1998 (Rel. 37, Created)
DE 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Arginine/ornithine antiporter.
GN ARCD OR CPE0170.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=13 / Type A;
RX MEDLINE=97199138; PubMed=9053381;
RA Ohtani K., Bando M., Swe T., Banu S., Oe M., Hayashi H., Shimizu T.;
RT "Collagenase gene (colA) is located in the 3'-flanking region of the
RL perfringolysin O (pfoA) locus in Clostridium perfringens.";
RL FEMS Microbiol. Lett. 146:155-159(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -!- FUNCTION: CATALYZES AN ELECTRONEUTRAL EXCHANGE BETWEEN ARGININE
CC AND ORNITHINE TO ALLOW HIGH-EFFICIENCY ENERGY CONVERSION IN THE
CC ARGinine DEIMINASE PATHWAY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
CC PERMEASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X97768; CAA66366.1; -.
DR EMBL; AP003185; BAB79876.1; -.
DR InterPro; IPR002293; AA/rel_prmease1.
DR InterPro; IPR004841; Permease.
DR InterPro; IPR004754; R/O_antiport.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00905; 2A0302; 1.
KW Transport; Antiport; Amino-acid transport; Transmembrane;
KW Complete proteome.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 83 103 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 130 150 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 203 223 POTENTIAL.
FT TRANSMEM 236 256 POTENTIAL.
FT TRANSMEM 284 304 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 359 379 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT TRANSMEM 456 476 POTENTIAL.
FT CONFLICT 83 83 MISSING (IN REF. 1).
FT CONFLICT 221 221 MISSING (IN REF. 1).
FT CONFLICT 352 352 A -> V (IN REF. 1).
FT CONFLICT 417 417 G -> E (IN REF. 1).
SQ SEQUENCE 478 AA; 51088 MW; A7BA24F7B53A344B CRC64;

Query Match 90.9%; Score 30; DB 1; Length 478;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWSMS 5
DB 96 GYWSLS 100

RESULT 8
ID PTSE STAXY STANDARD; PRT; 480 AA.
AC P51184;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE PTS system, sucrose-specific IIBC component (EIIBC-SCR) (Sucrose-
DE permease IIBC component) (Phosphotransferase enzyme II, BC component)
DE (EC 2.7.1.69) (EIIBC-SCR).
GN SCRA.
OS Staphylococcus xylosum.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1288;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DSM 20267 / Isolate C2A;
RX MEDLINE=94049686; PubMed=8232209;
RA Wagner E., Goetz F., Brueckner R.;
RT "Cloning and characterization of the scrA gene encoding the sucrose-
RT specific Enzyme II of the phosphotransferase system from
RT Staphylococcus xylosum.";
RL Mol. Gen. Genet. 241:33-41(1993).
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE IICD DOMAINS CONTAIN THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL; THE IIA DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR.
CC -!- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
CC histidine + sugar phosphate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
```



CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL; AF000062; BAA80640.1; --

CC HSP; P56206; IATI.

CC InterPro; IPR002106; AARNA ligaseII.

CC InterPro; IPR004154; HGTTP anticodon.

CC InterPro; IPR002314; tRNA-synt 2b.

CC InterPro; IPR002315; tRNA-synt\_gly.

CC Pfam; PF00587; tRNA-synt\_2b; 1.

CC Pfam; PF01129; HGTTP-anticodon; 1.

CC PRINTS; PRO1043; TRNASYNTHGLY.

CC TIGRFAMs; TIGR00389; glyS dimeric; 1.

CC PROSITE; PS00862; AA TRNA LIGASE II; 1.

CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

CC KW Complete proteome.

CC SQ SEQUENCE 583 AA; 66395 MW; 7D75DE7B5AA3P999 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 583;

Best Local Similarity 80.0%; Pred. No. 1.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWS 5

Db 277 GYWS 281

RESULT 10

CEA\_CITFR STANDARD; PRT; 592 AA.

ID CEA\_CITFR STANDARD; PRT; 592 AA.

AC P04480;

DT 13-AUG-1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Colicin A.

GN CAA.

OS Citrobacter freundii.

OG Plasmid ColA-CA31.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Citrobacter.

OX NCBI\_TaxID=546;

RN [1]

RP SEQUENCE FROM N.A.

RP MEDLINE=84036205; PubMed=6313941;

RA Morlon J., Chartier M., Bidaud M., Lazdunski C.;

RA "Complete nucleotide sequence of the structural gene for colicin A, a

RT gene translated at non-uniform rate.";

RL J. Mol. Biol. 170:271-285(1983).

RN [2]

RP SEQUENCE FROM N.A.

RP MEDLINE=88174422; PubMed=2832701;

RA Morlon J., Chartier M., Bidaud M., Lazdunski C.;

RT "The complete nucleotide sequence of the colicinogenic plasmid ColA.

RT High extent of homology with ColE1.";

RL Mol. Gen. Genet. 211:231-243(1988).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 389-592.

RP MEDLINE=92235820; PubMed=1373773;

RA Parker M.W., Postma J.P.M., Pattus F., Tucker A.D., Tsernoglou D.;

RA "Refined structure of the pore-forming domain of colicin A at 2.4-A

RT resolution.";

RL J. Mol. Biol. 224:639-657(1992).

CC -1- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF

CC TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING

CC TO DISSIPATION OF CELLULAR ENERGY.

CC -1- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE

CC AGAINST ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.

CC -1- SIMILARITY: BELONGS TO THE CHANNEL FORMING COLICIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -1- SIMILARITY: CONTAINS 1 PTS EIIB DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 PTS EIIB DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL; X69800; CAA49461.1; --

CC InterPro; IPR001996; PTS\_EIIB.

CC InterPro; IPR003352; PTS\_EIIC.

CC Pfam; PF00367; PTS\_EIIB; 1.

CC Pfam; PF02378; PTS\_EIIC; 1.

CC ProDom; PD001476; PTS\_EIIB; 1.

CC TIGRFAMs; TIGR00826; EIIBC; 1.

CC PROSITE; PS01035; PTS\_EIIB\_CYS; 1.

CC Phosphotransferase system; Sugar transport; Transferase;

CC KW Transmembrane; Inner membrane; Phosphorylation.

CC FT DOMAIN 1 ? EIIB DOMAIN.

FT MOD\_RES 26 26 PHOSPHORYLATION (BY SIMILARITY).

FT MOD\_RES 325 325 PHOSPHORYLATION (BY SIMILARITY).

FT SEQUENCE 480 AA; 51326 MW; AB4E1D9785D84E47 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 480;

Best Local Similarity 80.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWS 5

Db 290 GYWS 294

RESULT 9

SYG\_AERPE STANDARD; PRT; 583 AA.

ID SYG\_AERPE STANDARD; PRT; 583 AA.

AC Q9YBF8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glycyl-tRNA synthetase (EC 6.1.1.14) (Glycine-tRNA ligase) (GlyRS).

GN GLYS OR APE439.

OS Aeropyrum pernix.

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;

OC Desulfurococaceae; Aeropyrum.

OX NCBI\_TaxID=56636;

RN [1]

RP SEQUENCE FROM N.A.

RP STRAIN=K1;

RP MEDLINE=99310339; PubMed=10382966;

RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,

RA Jin N.K., Takahashi M., Sekine M., Baba S.-I., Ankaei A., Kosugi H.,

RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,

RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,

RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,

RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;

RA "Complete genome sequence of an aerobic hyper-thermophilic

RT Crenarchaeon, Aeropyrum pernix K1.";

RL DNA Res. 6:83-101(1999).

CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) = AMP + diphosphate

CC + glycyl-tRNA(Gly).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01008; AAA25503.1; -.
DR EMBL; M37402; AAA72879.1; -.
DR PIR; A03504; IKEBCA.
DR PDB; 1COL; 15-JUL-93.
DR InterPro; IPR000293; Channel_colicin.
DR Pfam; PF01024; Colicin; 1.
DR PRINTS; PR00280; Channel_colicin.
DR ProDom; PD002657; Channel_colicin; 1.
DR PROSITE; PS00276; CHANNEL_COLICIN; 1.
KW Antibiotic; Bacteriocin; Plasmid; Transmembrane; 3D-structure.
FT TRANSMEM 528 548
FT TRANSMEM 555 575
FT HELIX 396 418
FT HELIX 420 434
FT TURN 435 435
FT TURN 437 438
FT TURN 444 455
FT HELIX 456 456
FT TURN 458 460
FT HELIX 464 475
FT TURN 476 476
FT HELIX 479 489
FT HELIX 491 493
FT TURN 496 497
FT HELIX 498 515
FT HELIX 519 530
FT TURN 531 532
FT HELIX 535 552
FT HELIX 557 575
FT TURN 577 577
FT HELIX 578 586
FT TURN 587 588
SQ SEQUENCE 592 AA; 62992 MW; B80FA1F52A8CFC5D CRC64;

Query Match 90.9%; Score 30; DB 1; Length 592;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5
DB 136 GYWLS 140

RESULT 11
PFCK_NEOFR STANDARD; PRT; 608 AA.
AC P22130;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxykinase [GTP] [EC 4.1.1.32]
DE (Phosphoenolpyruvate carboxylase) (PEPCK).
OS Neocallimastix frontalis (Rumen fungus).
OC Eukaryota; Fungi; Chytridiomycota; Neocallimasticales;
OC Neocallimastixaceae; Neocallimastix.
OX NCBI_TaxID=4757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92184115; PubMed=1339359;
RA Raymond P., Geourjon C., Roub B., Durand R., Fevre M.;
RT "Sequence of the phosphoenolpyruvate carboxykinase-encoding cDNA from
RT the rumen anaerobic fungus Neocallimastix frontalis: comparison of
RT the amino acid sequence with animals and yeast.";
RL Gene 110:57-63 (1992).
CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]
CC FAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59372; AAA33553.1; -.
DR PIR; JQ1462; JQ1462.
DR InterPro; IPR000364; PEP_carboxykin.
DR Pfam; PF00821; PEPCK; 1.
DR ProDom; PD004738; PEP_carboxykin; 1.
DR PROSITE; PS00505; PEPCK_GTP; 1.
KW Gluconeogenesis; Lyase; Decarboxylase; GTP-binding.
FT NP_BIND 217 224 GTP (POTENTIAL).
FT ACT_SITE 269 269
SQ SEQUENCE 608 AA; 66904 MW; C4D2B249A92B7D26 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 608;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5
DB 478 GYWLS 482

RESULT 12
ENV_GALV STANDARD; PRT; 667 AA.
AC P21415;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein precursor [Contains: Knob protein GP70; Spike protein
DE P15E].
DE P15E].
GN ENV.
OS Gibbon ape leukemia virus.
OC Viruses; Retroviruses; Retroviridae; Gammaetrovirus.
OX NCBI_TaxID=11840;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90051069; PubMed=2683360;
RA Delassus S., Sonigo P., Wain-Hobson S.;
RT "Genetic organization of gibbon ape leukemia virus.";
RL Virology 173:205-213 (1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26927; AAA46811.1; -.
DR PIR; C32595; VCLJGL.
DR HSP; P03385; IMOF.
DR InterPro; IPR002050; Env_polyprotein.
DR Pfam; PF00429; ENV_polyprotein; 1.
KW Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal.
FT SIGNAL 1 42
FT CHAIN 43 489 KNOB PROTEIN GP70.
FT CHAIN 490 667 SPIKE PROTEIN P15E.
FT TRANSMEM 492 508 POTENTIAL.
FT TRANSMEM 637 653 POTENTIAL.
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

SQ SEQUENCE 667 AA; 73729 MW; 5FPB32F00D8631BE CRC64;  
 Query Match 90.9%; Score 30; DB 1; Length 667;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWS 5  
 DB 180 GYWS 184  
 RESULT 13  
 ID AGUA\_TIRE STANDARD; PRT; 847 AA.  
 AC Q99024;  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DE Alpha-glucuronidase precursor (EC 3.2.1.139) (Alpha-glucosiduronase) (GLXI).  
 DE Trichoderma reesei (Hypocrea jecorina).  
 OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreales; Hypocreaceae; Hypocrea.  
 ON NCBI\_TaxID=51453;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=QM9414 / Rut C-30;  
 RX MEDLINE=96257277; PubMed=8654984;  
 RA Margolles-Clark E., Saloheimo M., Siika-Aho M., Penttilae M.;  
 RT "The alpha-glucuronidase-encoding gene of Trichoderma reesei";  
 RL Gene 172:171-172(1996).  
 CC -!- FUNCTION: RELEASES 4-O-METHYLGLUCURONIC ACID FROM XYLAN.  
 CC -!- CATALYTIC ACTIVITY: An alpha-D-glucuronoside + H(2)O = an alcohol + D-glucuronate.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 67 OF GLYCOSYL HYDROLASES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC EMBL; Z68706; CAA92949.1; -  
 DR InterPro; IPR005154; Glyco\_hydro\_67.  
 DR Pfam; PF03648; Glyco\_hydro\_67; 1.  
 KW Xylan degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 847 ALPHA-GLUCURONIDASE.  
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 740 740 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 767 767 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 847 AA; 93424 MW; 214AAE487382FDB4 CRC64;  
 Query Match 90.9%; Score 30; DB 1; Length 847;  
 Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWS 5  
 DB 119 GYWS 123  
 RESULT 14  
 ID CAPP\_RHOMR STANDARD; PRT; 936 AA.

AC Q59757;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phosphoenolpyruvate carboxylase (EC 4.1.1.31) (PEPCASE) (PEPC).  
 GN PPC OR PEPC.  
 OS Rhodothermus marinus (Rhodothermus obamensis).  
 OC Bacteria; Bacteroidetes; Sphingobacteria; Sphingobacteriales;  
 OC Crenotrichaceae; Rhodothermus.  
 ON NCBI\_TaxID=29549;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OKD7;  
 RX MEDLINE=97420683; PubMed=9276668;  
 RA Takai K., Sako Y., Uchida A., Ishida Y.;  
 RT "Extremely thermostable phosphoenolpyruvate carboxylase from an extreme thermophile, Rhodothermus obamensis";  
 RL J. Biochem. 122:32-40(1997).  
 CC -!- FUNCTION: TO FORM OXALACETATE, A FOUR-CARBON DICARBOXYLIC ACID SOURCE FOR THE TRICARBOXYLIC ACID CYCLE. THE OPTIMUM TEMPERATURE FOR ACTIVITY IS 75 DEGREES CELSIUS. THE ENZYME EXHIBITS A PH OPTIMUM OF 8.0.  
 CC -!- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O + phosphoenolpyruvate + CO(2).  
 CC -!- COFACTOR: ABSOLUTE REQUIREMENT FOR DIVALENT CATIONS.  
 CC -!- PATHWAY: Tricarboxylic acid cycle.  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- SIMILARITY: BELONGS TO THE PEPCASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC EMBL; X99379; CAA67760.1; -  
 DR HSP; P00864; 1FY.  
 DR InterPro; IPR001449; PEPCase.  
 DR Pfam; PF00311; PEPCase; 1.  
 DR PRINTS; PR00150; PEPCARBLASE.  
 DR PROSITE; PS00393; PEPCASE\_2; 1.  
 DR PROSITE; PS00781; PEPCASE\_1; 1.  
 KW Lyase; Carbon dioxide fixation; Tricarboxylic acid cycle.  
 FT ACT\_SITE 155 155 BY SIMILARITY.  
 FT ACT\_SITE 595 595 BY SIMILARITY.  
 SQ SEQUENCE 936 AA; 107886 MW; C9AA94C3F26C345A CRC64;  
 Query Match 90.9%; Score 30; DB 1; Length 936;  
 Best Local Similarity 80.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWS 5  
 DB 598 GYWS 602  
 RESULT 15  
 ID CZCA\_ALCEU STANDARD; PRT; 1063 AA.  
 AC P13511; P94142;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cobalt-zinc-cadmium resistance protein Czca (Cation efflux system protein czca).  
 DE protein czca).  
 GN CZCA.  
 OS Alkaligenes eutrophus (Ralstonia eutropha).  
 OG Plasmid pMOL30.  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 ON NCBI\_TaxID=510;

RN SEQUENCE FROM N.A.  
 RP STRAIN=CH34;  
 RC MEDLINE=90017477; PubMed=2678100;  
 RX Nies D.H., Nies A., Chu L., Silver S.;  
 RA "Expression and nucleotide sequence of a plasmid-determined divalent  
 RT cation efflux system from *Alcaligenes eutrophus*";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7351-7355(1989).  
 RN [2]  
 RP REVISIONS TO 918 AND 959-960.  
 RC STRAIN=CH34;  
 RA van der Lelie D., Schwuchow T., Wuerztz S., Schwidetzky U.,  
 RA Baeyens W., Scheel P.O., Nies D.H.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: HAS A LOW CATION TRANSPORT ACTIVITY FOR CO(2+), IT IS  
 CC ESSENTIAL FOR THE EXPRESSION OF COBALT, ZINC, AND CADMIUM  
 CC RESISTANCE. CZCA AND CZCB TOGETHER WOULD ACT IN ZN(2+) EFFLUX  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLY).  
 CC -!- INDUCTION: BY CADMIUM, COPPER AND ZINC.  
 CC -!- SIMILARITY: BELONGS TO THE ACRB/ACRD/ACRF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; X98451; CAA67084.1; -;  
 DR FIR; A33830; A33830.  
 DR InterPro; IPR001036; Acrflvln\_res.  
 DR InterPro; IPR004763; CzCA.  
 DR Pfam; PF00873; ACR\_tran; 1.  
 DR TIGRFAMs; TIGR00914; 2A0601; 1.  
 KW Plasmid; Transport; Zinc; Cobalt; Cadmium resistance; Transmembrane.  
 FT TRANSMEM 14 29 POTENTIAL.  
 FT TRANSMEM 350 370 POTENTIAL.  
 FT TRANSMEM 452 472 POTENTIAL.  
 FT TRANSMEM 487 507 POTENTIAL.  
 FT TRANSMEM 534 554 POTENTIAL.  
 FT TRANSMEM 883 903 POTENTIAL.  
 FT TRANSMEM 906 926 POTENTIAL.  
 FT TRANSMEM 937 957 POTENTIAL.  
 FT TRANSMEM 982 1004 POTENTIAL.  
 FT TRANSMEM 1013 1033 POTENTIAL.  
 SQ SEQUENCE 1063 AA; 115668 MW; 1E95B7923996AEBC CRC64;

Qy 1 GYVMS 5  
 Db 862 GYVMT 866

Search completed: April 22, 2003, 12:51:50  
 Job time: 4.5 secs

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:47:43 ; Search time 10 Seconds  
(without alignments)  
103.024 Million cell updates/sec

Title: US-09-674-716B-9

Perfect score: 33

Sequence: 1 GYMS 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	67	Q9HN48	Q9HN48 halobacteri
2	33	100.0	377	Q9I867	Q9I867 campyloform
3	33	100.0	377	Q9IAH6	Q9IAH6 boulengerom
4	33	100.0	377	Q9IAH4	Q9IAH4 brienyomyrus
5	33	100.0	377	Q9IAH3	Q9IAH3 brienyomyrus
6	33	100.0	377	Q9IAH2	Q9IAH2 brienyomyrus
7	33	100.0	377	Q9IAH1	Q9IAH1 brienyomyrus
8	33	100.0	377	Q9IAH0	Q9IAH0 campyloform
9	33	100.0	377	Q9IAG9	Q9IAG9 campyloform
10	33	100.0	377	Q9IAG7	Q9IAG7 genyomyrus
11	33	100.0	377	Q9IAG6	Q9IAG6 enathionemus
12	33	100.0	377	Q9IAG4	Q9IAG4 hippopotamy
13	33	100.0	377	Q9IAG3	Q9IAG3 hippopotamy
14	33	100.0	377	Q9IAG2	Q9IAG2 hippopotamy
15	33	100.0	377	Q9IAG1	Q9IAG1 hyperopisus
16	33	100.0	377	Q9IAF9	Q9IAF9 ivindomyrus

17	33	100.0	377	13	Q9IAP8	Q9Iaf8 marcusenius
18	33	100.0	377	13	Q9IAE7	Q9Iaf7 marcusenius
19	33	100.0	377	13	Q9IAF6	Q9Iaf6 marcusenius
20	33	100.0	377	13	Q9IAF5	Q9Iaf5 marcusenius
21	33	100.0	377	13	Q9IAF4	Q9Iaf4 marcusenius
22	33	100.0	377	13	Q9IAF0	Q9Iaf0 mormyru ov
23	33	100.0	377	13	Q9IAE9	Q9Iae9 mormyru ru
24	33	100.0	377	13	Q9IAE6	Q9Iae6 paramormyro
25	33	100.0	377	13	Q9IAE1	Q9Iae1 pollimyrus
26	33	100.0	377	13	Q9IAE0	Q9Iae0 pollimyrus
27	33	100.0	377	13	Q9IAD9	Q9Iad9 stomatorhin
28	33	100.0	377	13	Q9IAD8	Q9Iad8 stomatorhin
29	33	100.0	377	13	Q9IAD7	Q9Iad7 stomatorhin
30	33	100.0	377	13	Q9IAD6	Q9Iad6 stomatorhin
31	33	100.0	377	13	Q9IAD5	Q9Iad5 stomatorhin
32	33	100.0	721	10	Q9FRN4	Q9frn4 oryza sativ
33	33	100.0	1028	5	Q8SXA6	Q8sxa6 drosophila
34	33	100.0	1043	16	Q97E41	Q97e41 clostridium
35	33	100.0	1254	5	Q9VF51	Q9vf51 drosophila
36	30	90.9	77	2	Q9SAW3	Q9saW3 escherichia
37	30	90.9	79	2	Q9WTF7	Q9wtf7 escherichia
38	30	90.9	79	16	Q9JMR3	Q9jmr3 escherichia
39	30	90.9	86	2	Q9ZGR9	Q9zgr9 escherichia
40	30	90.9	175	10	Q8XS1	Q8xsl arabidopsis
41	30	90.9	253	2	Q32816	Q32816 lactococcus
42	30	90.9	272	8	Q9G7X3	Q9g7x3 philotrypes
43	30	90.9	285	17	Q8U3V4	Q8u3v4 pyrococcus
44	30	90.9	298	16	Q925X5	Q925x5 listeria in
45	30	90.9	314	4	Q9BTL6	Q9btl6 homo sapien

ALIGNMENTS

RESULT 1

Q9HN48	PRELIMINARY;	PRT;	67 AA.
ID	Q9HN48		
AC	Q9HN48;		
DT	01-MAR-2001 (TRENBLrel. 16, Created)		
DT	01-MAR-2001 (TRENBLrel. 16, Last sequence update)		
DT	01-JUN-2001 (TRENBLrel. 17, Last annotation update)		
DE	Vng2253h.		
GN	VNG2253H.		
OS	Halobacterium sp. (strain NRC-1).		
OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;		
OC	Halobacteriaceae; Halobacterium.		
OX	NCBI_TaxID=64091;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20504483; PubMed=11016950;		
RA	Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,		
RA	Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,		
RA	Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,		
RA	Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,		
RA	Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,		
RA	Isebnarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,		
RA	Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,		
RA	Edhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;		
RL	"Genome sequence of Halobacterium species NRC-1.;"		
DR	EMBL; A8005110; AAG20373.1; -.		
KW	Complete proteome.		
SQ	SEQUENCE 67 AA; 6601 MW; A6BAF23FCBB8B648 CRC64;		

Query Match 100.0%; Score 33; DB 17; Length 67;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYMS 5

Db 59 GYMS 63

```
RESULT 2
Q91867
ID Q91867 PRELIMINARY; PRT; 377 AA.
AC Q91867;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Recombination-activating protein 2 (Fragment).
OS Campylomorphomys tamandua.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Campylomorphomys.
OX NCBI_TaxID=91719;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20115608; PubMed=10648209;
RA Sullivan J.P., Lavoue S., Hopkins C.D.;
RT "Molecular systematics of the African electric fishes (Mormyroidea;
RT Teleostei) and a model for the evolution of their electric organs.";
RL J. Exp. Biol. 203:665-683(2000).
DR EMBL; AF201625; AAF43336.1; -.
DR EMBL; AF201624; AAF43335.1; -.
DR InterPro; IPR004321; RAG2.
DR Pfam; PF03089; RAG2; 1.
FT NON_TER 1
FT NON_TER 377
SQ SEQUENCE 377 AA; 41387 MW; D52A9E361A56AB43 CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
DB 368 GYWS 372

RESULT 3
Q91A86
ID Q91A86 PRELIMINARY; PRT; 377 AA.
AC Q91A86;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Recombination-activating protein 2 (Fragment).
OS Boulengeromorphomys knoepfieri.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Boulengeromorphomys.
OX NCBI_TaxID=91717;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20115608; PubMed=10648209;
RA Sullivan J.P., Lavoue S., Hopkins C.D.;
RT "Molecular systematics of the African electric fishes (Mormyroidea;
RT Teleostei) and a model for the evolution of their electric organs.";
RL J. Exp. Biol. 203:665-683(2000).
DR EMBL; AF201616; AAF43327.1; -.
DR InterPro; IPR004321; RAG2.
DR Pfam; PF03089; RAG2; 1.
FT NON_TER 1
FT NON_TER 377
SQ SEQUENCE 377 AA; 41428 MW; B60EDE613EA0FDBE CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
DB 368 GYWS 372
```

```
RESULT 4
Q91A84
ID Q91A84 PRELIMINARY; PRT; 377 AA.
AC Q91A84;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Recombination-activating protein 2 (Fragment).
OS Brienomorphomys hopkinsi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Brienomorphomys.
OX NCBI_TaxID=112141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20115608; PubMed=10648209;
RA Sullivan J.P., Lavoue S., Hopkins C.D.;
RT "Molecular systematics of the African electric fishes (Mormyroidea;
RT Teleostei) and a model for the evolution of their electric organs.";
RL J. Exp. Biol. 203:665-683(2000).
DR EMBL; AF201618; AAF43329.1; -.
DR InterPro; IPR004321; RAG2.
DR Pfam; PF03089; RAG2; 1.
FT NON_TER 1
FT NON_TER 377
SQ SEQUENCE 377 AA; 41403 MW; 0A4599C6604C8123 CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
DB 368 GYWS 372

RESULT 5
Q91A83
ID Q91A83 PRELIMINARY; PRT; 377 AA.
AC Q91A83;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Recombination-activating protein 2 (Fragment).
OS Brienomorphomys longicaudatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Brienomorphomys.
OX NCBI_TaxID=112142;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20115608; PubMed=10648209;
RA Sullivan J.P., Lavoue S., Hopkins C.D.;
RT "Molecular systematics of the African electric fishes (Mormyroidea;
RT Teleostei) and a model for the evolution of their electric organs.";
RL J. Exp. Biol. 203:665-683(2000).
DR EMBL; AF201619; AAF43330.1; -.
DR InterPro; IPR004321; RAG2.
DR Pfam; PF03089; RAG2; 1.
FT NON_TER 1
FT NON_TER 377
SQ SEQUENCE 377 AA; 41403 MW; 0A4599C6604C8123 CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
DB 368 GYWS 372

RESULT 6
```

```

Q9IAH2
ID Q9IAH2 PRELIMINARY; PRT; 377 AA.
AC Q9IAH2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Recombination-activating protein 2 (Fragment).
OS Brienomyrus niger.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Brienomyrus.
OX NCBI_TaxID=42637;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20115608; PubMed=10648209;
RA Sullivan J.P., Lavoue S., Hopkins C.D.;
RT "Molecular systematics of the African electric fishes (Mormyroidea;
RT Teleostei) and a model for the evolution of their electric organs.";
RL J. Exp. Biol. 203:665-683(2000).
DR EMBL; AF201620; AAF43331.1; -.
DR InterPro; IPR004321; RAG2.
DR Pfam; PF03089; RAG2; 1.
FT NON_TER 1
FT NON_TER 377
SQ SEQUENCE 377 AA; 41522 MW; 2E93DC79A8B6EC4A CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
DB 368 GYWS 372

RESULT 7
Q9IAH1
ID Q9IAH1 PRELIMINARY; PRT; 377 AA.
AC Q9IAH1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Recombination-activating protein 2 (Fragment).
OS Brienomyrus sp. CU79740.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Brienomyrus.
OX NCBI_TaxID=112278;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20115608; PubMed=10648209;
RA Sullivan J.P., Lavoue S., Hopkins C.D.;
RT "Molecular systematics of the African electric fishes (Mormyroidea;
RT Teleostei) and a model for the evolution of their electric organs.";
RL J. Exp. Biol. 203:665-683(2000).
DR EMBL; AF201621; AAF43332.1; -.
DR InterPro; IPR004321; RAG2.
DR Pfam; PF03089; RAG2; 1.
FT NON_TER 1
FT NON_TER 377
SQ SEQUENCE 377 AA; 41475 MW; 735853EEA67408FE CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
DB 368 GYWS 372

RESULT 8
Q9IAH0
ID Q9IAH0 PRELIMINARY; PRT; 377 AA.
AC Q9IAH0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Recombination-activating protein 2 (Fragment).
OS Campylomormyrus numenius.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Campylomormyrus.
OX NCBI_TaxID=112143;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20115608; PubMed=10648209;
RA Sullivan J.P., Lavoue S., Hopkins C.D.;
RT "Molecular systematics of the African electric fishes (Mormyroidea;
RT Teleostei) and a model for the evolution of their electric organs.";
RL J. Exp. Biol. 203:665-683(2000).
DR EMBL; AF201622; AAF43333.1; -.
DR InterPro; IPR004321; RAG2.
DR Pfam; PF03089; RAG2; 1.
FT NON_TER 1
FT NON_TER 377
SQ SEQUENCE 377 AA; 41360 MW; C0E574061A4C986A CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
DB 368 GYWS 372

RESULT 9
Q9IAG9
ID Q9IAG9 PRELIMINARY; PRT; 377 AA.
AC Q9IAG9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Recombination-activating protein 2 (Fragment).
OS Campylomormyrus sp. JPS1-2000.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Campylomormyrus.
OX NCBI_TaxID=112279;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20115608; PubMed=10648209;
RA Sullivan J.P., Lavoue S., Hopkins C.D.;
RT "Molecular systematics of the African electric fishes (Mormyroidea;
RT Teleostei) and a model for the evolution of their electric organs.";
RL J. Exp. Biol. 203:665-683(2000).
DR EMBL; AF201623; AAF43334.1; -.
DR InterPro; IPR004321; RAG2.
DR Pfam; PF03089; RAG2; 1.
FT NON_TER 1
FT NON_TER 377
SQ SEQUENCE 377 AA; 41357 MW; DDA7ED06162C799A CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
DB 368 GYWS 372

RESULT 10
Q9IAG7
ID Q9IAG7 PRELIMINARY; PRT; 377 AA.

```

AC Q91AG7;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Recombination-activating protein 2 (Fragment).  
 OS Genomyrus donny.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Mormyridae; Genomyrus.  
 OX NCBI\_TaxID=89589;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20115608; PubMed=10648209;  
 RA Sullivan J.P., Lavoue S., Hopkins C.D.;  
 RT "Molecular systematics of the African electric fishes (Mormyroidea: Teleostei) and a model for the evolution of their electric organs.";  
 RL J. Exp. Biol. 203:665-683(2000).  
 DR EMBL; AF201627; AAF43338.1; -.  
 DR InterPro; IPR004321; RAG2.  
 DR Pfam; PF03089; RAG2; 1.  
 FT NON\_TER 1  
 FT NON\_TER 377  
 SQ SEQUENCE 377 AA; 41287 MW; F43A9AFC686C0F67 CRC64;  
 Query Match 100.0%; Score 33; DB 13; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWMS 5  
 Db 368 GYWMS 372  
 RESULT 11  
 Q91AG6  
 ID Q91AG6 PRELIMINARY; PRT; 377 AA.  
 AC Q91AG6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Recombination-activating protein 2 (Fragment).  
 OS Gnathonemus petersii.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Mormyridae; Gnathonemus.  
 OX NCBI\_TaxID=42645;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20115608; PubMed=10648209;  
 RA Sullivan J.P., Lavoue S., Hopkins C.D.;  
 RT "Molecular systematics of the African electric fishes (Mormyroidea: Teleostei) and a model for the evolution of their electric organs.";  
 RL J. Exp. Biol. 203:665-683(2000).  
 DR EMBL; AF201628; AAF43339.1; -.  
 DR InterPro; IPR004321; RAG2.  
 DR Pfam; PF03089; RAG2; 1.  
 FT NON\_TER 1  
 FT NON\_TER 377  
 SQ SEQUENCE 377 AA; 41372 MW; 2406A9E13FE299F CRC64;  
 Query Match 100.0%; Score 33; DB 13; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWMS 5  
 Db 368 GYWMS 372  
 RESULT 12  
 Q91AG4  
 ID Q91AG4 PRELIMINARY; PRT; 377 AA.  
 AC Q91AG4;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Recombination-activating protein 2 (Fragment).  
 OS Hippopotamys discorhynchus (Zambezi parrotfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Mormyridae; Hippopotamyrus.  
 OX NCBI\_TaxID=112147;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20115608; PubMed=10648209;  
 RA Sullivan J.P., Lavoue S., Hopkins C.D.;  
 RT "Molecular systematics of the African electric fishes (Mormyroidea: Teleostei) and a model for the evolution of their electric organs.";  
 RL J. Exp. Biol. 203:665-683(2000).  
 DR EMBL; AF201630; AAF43341.1; -.  
 DR InterPro; IPR004321; RAG2.  
 DR Pfam; PF03089; RAG2; 1.  
 FT NON\_TER 1  
 FT NON\_TER 377  
 SQ SEQUENCE 377 AA; 41293 MW; 9F7C51AF898736C4 CRC64;  
 Query Match 100.0%; Score 33; DB 13; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWMS 5  
 Db 368 GYWMS 372  
 RESULT 13  
 Q91AG3  
 ID Q91AG3 PRELIMINARY; PRT; 377 AA.  
 AC Q91AG3;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Recombination-activating protein 2 (Fragment).  
 OS Hippopotamyrus pictus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Mormyridae; Hippopotamyrus.  
 OX NCBI\_TaxID=112148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20115608; PubMed=10648209;  
 RA Sullivan J.P., Lavoue S., Hopkins C.D.;  
 RT "Molecular systematics of the African electric fishes (Mormyroidea: Teleostei) and a model for the evolution of their electric organs.";  
 RL J. Exp. Biol. 203:665-683(2000).  
 DR EMBL; AF201631; AAF43342.1; -.  
 DR InterPro; IPR004321; RAG2.  
 DR Pfam; PF03089; RAG2; 1.  
 FT NON\_TER 1  
 FT NON\_TER 377  
 SQ SEQUENCE 377 AA; 41491 MW; 1BE26037590CA937 CRC64;  
 Query Match 100.0%; Score 33; DB 13; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWMS 5  
 Db 368 GYWMS 372  
 RESULT 14  
 Q91AG2  
 ID Q91AG2 PRELIMINARY; PRT; 377 AA.  
 AC Q91AG2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)



DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Recombination-activating protein 2 (Fragment).  
 OS Hippopotamyrus wilverthi.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Mormyridae; Hippopotamyrus.  
 OX NCBI\_TaxID=112149;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20115608; PubMed=10648209;  
 RA Sullivan J.P., Lavoue S., Hopkins C.D.;  
 RT "Molecular systematics of the African electric fishes (Mormyroidea;  
 RT Teleostei) and a model for the evolution of their electric organs.";  
 RL J. Exp. Biol. 203:665-683 (2000).  
 DR EMBL; AF201632; AAF43343.1; -.  
 DR InterPro; IPR004321; RAG2.  
 DR Pfam; PF03089; RAG2; 1.  
 FT NON\_TER 1  
 FT NON\_TER 377 377  
 SQ SEQUENCE 377 AA; 41285 MW; C62A314A621D3AEE CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWMS 5  
 DB 368 GYWMS 372

RESULT 15  
 Q9IAG1 PRELIMINARY; PRT; 377 AA.  
 AC Q9IAG1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Recombination-activating protein 2 (Fragment).  
 OS Hyperopisus bebe.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Mormyridae; Hyperopisus.  
 OX NCBI\_TaxID=91725;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20115608; PubMed=10648209;  
 RA Sullivan J.P., Lavoue S., Hopkins C.D.;  
 RT "Molecular systematics of the African electric fishes (Mormyroidea;  
 RT Teleostei) and a model for the evolution of their electric organs.";  
 RL J. Exp. Biol. 203:665-683 (2000).  
 DR EMBL; AF201633; AAF43344.1; -.  
 DR InterPro; IPR004321; RAG2.  
 DR Pfam; PF03089; RAG2; 1.  
 FT NON\_TER 1  
 FT NON\_TER 377 377  
 SQ SEQUENCE 377 AA; 41324 MW; C35A2BBE34EF6FC CRC64;

Query Match 100.0%; Score 33; DB 13; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWMS 5  
 DB 368 GYWMS 372

Search completed: April 22, 2003, 12:53:50  
 Job time : 11 secs



GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:45:19 ; Search time 12.6786 Seconds  
(without alignments)  
52.550 Million cell updates/sec

Title: US-09-674-716B-9

Perfect score: 33

Sequence: 1 GYWS 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq 101002:\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	5	21	Light chain CDR H1
2	33	100.0	137	21	Mouse anti-CD23 MA
3	33	100.0	444	21	Humanised anti-CD2
4	33	100.0	1254	22	Drosophila melanog
5	30	90.9	40	22	Novel human diagno
6	30	90.9	43	17	V39, monoclonal an
7	30	90.9	64	20	Human secreted pro
8	30	90.9	181	20	Chlamydia pneumoni
9	30	90.9	210	22	Novel signal trans
10	30	90.9	242	21	Rat HT glycoprotei

11	30	90.9	242	22	AA56053	Skin cell protein,
12	30	90.9	242	23	ABB72253	Rat protein isolat
13	30	90.9	242	23	ABB72290	Rat protein isolat
14	30	90.9	242	23	ABB72301	Rat protein isolat
15	30	90.9	243	23	ABP39269	Staphylococcus epi
16	30	90.9	310	22	AAU35317	Enterococcus faeca
17	30	90.9	310	22	AAU35317	GALV retrovirus su
18	30	90.9	315	23	ABB55045	Lactococcus lactis
19	30	90.9	359	21	AAU31982	Arabidopsis thalia
20	30	90.9	374	22	AAU31982	Activating polyep
21	30	90.9	379	21	AAU31981	Arabidopsis thalia
22	30	90.9	392	21	AAU31981	Human secreted pro
23	30	90.9	420	20	AAU31981	Amino acid sequenc
24	30	90.9	420	20	AAU31981	EGF-like homologue
25	30	90.9	420	20	AAU31981	Human PRO214 prote
26	30	90.9	420	21	AAU31981	Human PRO214 amino
27	30	90.9	420	21	AAU31981	Human polypeptide
28	30	90.9	420	22	AAU31981	Human PRO214 polyp
29	30	90.9	420	22	AAU31981	PRO214. Homo sapi
30	30	90.9	420	22	AAU31981	Human PRO214 prote
31	30	90.9	420	22	AAU31981	Mouse TANGO 206 po
32	30	90.9	420	22	AAU31981	Human TANGO 206 va
33	30	90.9	420	22	AAU31981	Human TANGO 206 va
34	30	90.9	420	22	AAU31981	Human TANGO 206 va
35	30	90.9	420	22	AAU31981	Human TANGO 206 va
36	30	90.9	420	22	AAU31981	Mouse TANGO 206 va
37	30	90.9	420	22	AAU31981	Mouse TANGO 206 va
38	30	90.9	420	22	AAU31981	Mouse TANGO 206 va
39	30	90.9	420	22	AAU31981	Human EXMAD-6 SQ
40	30	90.9	420	23	AAU31981	Human angiogenesis
41	30	90.9	420	23	AAU31981	Human PRO214 prote
42	30	90.9	420	23	AAU31981	Human PRO protein,
43	30	90.9	448	21	AAU31981	Streptococcus pneu
44	30	90.9	449	21	AAU31981	Streptococcus pneu
45	30	90.9	451	23	ABP25569	Streptococcus poly

## ALIGNMENTS

RESULT 1

AAU32257

ID AAU32257 standard; Peptide; 5 AA.

XX AAU32257;

AC AAU32257;

XX

DT 15-FEB-2000 (first entry)

DE Light chain CDR H1 of mouse anti-CD23 MAB C11.

XX

KW CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse;

KW monoclonal antibody; chimeric antibody; humanised antibody;

KW complementarity determining region; CDR; autoimmune disease;

KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;

KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;

KW urticaria; nephrotic syndrome; glomerulonephritis;

KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;

KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;

KW graft-versus-host disease; COPD; bronchitis; diabetes;

KW B-cell malignancy; therapy.

XX Mus musculus.

OS

XX WO9958679-A1.

XX

PD 18-NOV-1999.

XX

PF 07-MAY-1999; 99WO-GB01434.

XX

PR 09-MAY-1998; 98GB-0009839.

XX

PA (GLAX ) GLAXO GROUP LTD.

XX



FH Key Location/Qualifiers  
 FT Region 1..30  
 FT /note= "framework region 1"  
 FT Region 31..35  
 FT /note= "CDR 1"  
 FT Region 36..49  
 FT /note= "framework region 2"  
 FT Region 50..68  
 FT /note= "CDR 2"  
 FT Region 69..100  
 FT /note= "framework region 3"  
 FT Region 101..103  
 FT /note= "CDR 3"  
 FT Region 104..111  
 FT /note= "framework region 4"  
 FT Region 112..444  
 FT /note= "constant region"  
 FT  
 FT

PN WO9958679-A1.

XX 18-NOV-1999.

PP 07-MAY-1999; 99WO-GB01434.

PR 09-MAY-1998; 98GB-0009839.

XX (GLAX ) GLAXO GROUP LTD.

PI Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;

DR WPI: 2000-053101/04.

DR N-PSDB; AAZ34748.

XX Cell receptor specific antibodies useful for treating e.g. arthritis, diabetes, multiple sclerosis and psoriasis -

PS Claim 9; Fig 4; 81pp; English.

CC This amino acid sequence represents the heavy chain of humanised  
 CC anti-CD23 (FCER1) monoclonal antibody C11, composed of a human  
 CC framework (HSGKRVII) and the heavy chain complementarity  
 CC determining regions (see AY32257-59) of murine antibody C11. The  
 CC DNA was constructed by splice overlap PCR. The invention provides  
 CC altered antibodies, such as chimeric or humanised antibodies, which  
 CC comprise sufficient of the amino acid sequences of the C11 light  
 CC and heavy chain complementarity determining regions to render them  
 CC capable of binding to the CD23 type II molecule expressed on  
 CC haematopoietic cells. The antibodies are used to block soluble  
 CC CD23 formation in human therapy, for the treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.

XX Sequence 444 AA;

Query Match 100.0%; Score 33; DB 21; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GYWS 5

DB 31 GYWS 35

RESULT 4

ABB62006

ID ABB62006 standard; Protein; 1254 AA.  
 AC ABB62006;  
 DT 26-MAR-2002 (first entry)  
 XX Drosophila melanogaster polypeptide SEQ ID NO 12810.  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical.  
 XX Drosophila melanogaster.  
 OS WO200171042-A2.  
 XX 27-SEP-2001.  
 PD 23-MAR-2001; 2001WO-US09231.  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX (PEXE ) PE CORP NY.  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI: 2001-656860/75.  
 DR N-PSDB; ABL06109.  
 DR New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX Disclosure; SEQ ID NO 12810; 21pp + Sequence Listing; English.  
 PS The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC\* (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 1254 AA;  
 SQ Query Match 100.0%; Score 33; DB 22; Length 1254;  
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 GYWS 5  
 DB 1139 GYWS 1143  
 RESULT 5  
 ID ABB62006 standard; Protein; 1254 AA.  
 AC ABB62006;  
 DT 13-FEB-2002 (first entry)  
 XX Novel human diagnostic protein #7524.  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KM food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.

PN WO200175067-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US08631.  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 XX N-PSDB; AAS71720.  
 DR  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 XX Claim 20; SEQ ID No 37892; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAG0010-AAG3037 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 40 AA;  
 Query Match 90.9%; Score 30; DB 22; Length 40;  
 Best Local Similarity 80.0%; Pred. No. 71;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWS 5  
 DB 33 GYWS 37  
 RESULT 6  
 AAR95472  
 ID AAR95472 standard; peptide; 43 AA. ✓  
 XX AAR95472;  
 AC  
 XX 05-NOV-1996 (first entry)  
 DT V39, monoclonal antibody SM-3 derived antigen binding peptide.  
 DE  
 XX Abtide; prostate specific mucin antigen; human prostate cancer; LNCAp;  
 KW diagnostic; detection; imaging; tumour; phage; peptide library; breast;  
 KW polymorphic; epithelial.  
 XX Synthetic.  
 OS  
 XX WO9609411-A1.  
 PN

XX 28-MAR-1996.  
 PD  
 XX 20-SEP-1995; 95WO-US11934.  
 PF  
 XX 07-JUN-1995; 95US-0488161.  
 PR 21-SEP-1994; 94US-0310192.  
 XX (CYTO-) CYTOGEN CORP.  
 PA Alvarez VL;  
 XX WPI; 1996-188471/19.  
 DR  
 XX New isolated peptide(s) with specific binding activities - obtd. by  
 PT screening random peptide libraries, for use in diagnostic and  
 PT therapeutic comps.  
 PT  
 XX Claim 38; Page 75; 106pp; English.  
 PS  
 XX AAR95459-R95509 are antigen binding peptides ("abtides") derived from  
 CC the monoclonal antibody SM-3 which recognises a specific polymorphic  
 CC epithelial mucin tumour antigen found on human breast cancer cells.  
 CC The abtides are identified from random peptide libraries using  
 CC specific ligand binding. Abtides mimic the binding specificity of  
 CC large molecules such as antibodies and receptors but have a much  
 CC smaller size allowing their production at a lower cost and reducing  
 CC the extent of their immunogenicity aiding in vivo delivery. The  
 CC abtides are useful for the diagnosis, detection, imaging and  
 CC treatment of disease, e.g. tumours, prostate cancer and breast  
 CC cancer.  
 XX  
 SQ Sequence 43 AA;  
 Query Match 90.9%; Score 30; DB 17; Length 43;  
 Best Local Similarity 80.0%; Pred. No. 77;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWS 5  
 DB 6 GYWS 10  
 RESULT 7  
 AAY07813  
 ID AAY07813 standard; Protein; 64 AA.  
 XX AAY07813;  
 AC  
 XX 06-JUL-1999 (first entry)  
 DT Human secreted protein fragment encoded from gene 70.  
 XX  
 DE Human; secreted protein; prevention; treatment; protein therapy;  
 KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;  
 KW developmental abnormality; foetal deficiency; blood disorder; lymphoma;  
 KW leukemia; immune system disorder; autoimmune disease; hepatic disease;  
 KW renal disease; inflammation; allergy; asthma; sepsis; diabetes; AIDS;  
 KW Alzheimer's disease; cognitive disorder; schizophrenia; osteoporosis;  
 KW arthritis; psoriasis; digestive; endocrine; infection.  
 XX Homo sapiens.  
 OS  
 XX WO9909155-A1.  
 PN  
 XX 25-FEB-1999.  
 PD  
 XX 18-AUG-1998; 98WO-US17044.  
 PF  
 XX 16-JUN-1998; 98US-0092956.  
 PR 15-JUL-1998; 98US-0092956.  
 PR 19-AUG-1997; 97US-0056368.  
 PR 19-AUG-1997; 97US-0056368.  
 PR

```

PR 19-AUG-1997; 97US-0056535.
PR 19-AUG-1997; 97US-0056555.
PR 19-AUG-1997; 97US-0056556.
PR 19-AUG-1997; 97US-0056628.
PR 19-AUG-1997; 97US-0056629.
PR 19-AUG-1997; 97US-0056726.
PR 19-AUG-1997; 97US-0056728.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer LA, Duan R, Ebner R, Endress GA, Feng P;
PI Florence C, Florence KA, Komatsoulis GA, Lafleur DW;
PI Moore PA, Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR;
PI Young PE;
XX
XX WPI; 1999-190160/16.
XX N-PSDB; AAX37438.
XX
XX New isolated human genes and the secreted polypeptides they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
XX Claim 1b; Page 264; 280pp; English.
XX
XX This invention describes novel isolated human secreted proteins and
XX their encoding nucleic acid sequences. The products of the invention
XX are useful for preventing, treating or ameliorating medical conditions
XX e.g. by protein or gene therapy. Also pathological conditions can be
XX diagnosed by determining the presence or amount of expression of
XX the new polypeptides in a sample or by determining the presence or
XX absence of mutations in the new polynucleotides. Specific uses are
XX described for each of the 70 polynucleotides, based on which tissues
XX they are most highly expressed in, and include developing products for
XX the diagnosis or treatment of cancer, tumours, neurodegenerative
XX disorders, developmental abnormalities and foetal deficiencies, blood
XX disorders, leukemias, diseases of the immune system, autoimmune
XX diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
XX asthma, sepsis, diabetes, Alzheimer's and cognitive disorders,
XX schizophrenia, osteoporosis, arthritis, psoriasis, digestive/endocrine
XX disorders, infections and AIDS. The human secreted proteins of the
XX invention are represented in AAX07744-Y07850 and the encoding nucleic
XX acids are represented in AAX37369-X37441.
XX
XX Sequence 64 AA;
XX
XX Query Match 90.9%; Score 30; DB 20; Length 64;
XX Best Local Similarity 80.0%; Pred. No. 1.2e+02; Mismatches 0; Gaps 0;
XX Matches 4; Conservative 1; Indels 0;
XX
XX QY 1 GYXMS 5
XX |||:|
XX Db 45 GYXLS 49
XX
XX RESULT 8
XX AAY35482
XX ID AAY35482 standard; Protein; 181 AA.
XX AC AAY35482;
XX
XX 13-SEP-1999 (first entry)
XX
XX Chlamydia pneumoniae transmembrane protein sequence.
XX
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX vaccine; neutralising epitope.
XX
XX Chlamydia pneumoniae.
XX OS
XX WO927105-A2.
XX
XX 03-JUN-1999.
XX

```

```

XX 20-NOV-1998; 98WO-IB01890.
XX
XX 04-NOV-1998; 98US-0107078.
XX 21-NOV-1997; 97FR-0014673.
XX (GEST ) GENSET.
XX
XX Griffais R;
XX
XX WPI; 1999-357842/30.
XX
XX Genome sequence of Chlamydia pneumoniae
XX
XX Page 1244; Disclosure; 1912pp; English.
XX
XX AAY34584-V35879 represent the proteins encoded by all the open reading
XX frames in the complete genome (see AAX31990) of Chlamydia pneumoniae.
XX C. pneumoniae causes respiratory disease such as pneumonia and
XX bronchitis and is thought to be a contributing factor in heart
XX disease, sarcoidosis, sinusitis, purulent otitis media, erythema
XX nodosum or pharyngitis. The polypeptides encoded by the open reading
XX frames of the C. pneumoniae genome (see AAY34584-V35879) can be used in
XX immunogenic compositions as vaccines. Vectors containing C. pneumoniae
XX nucleotide sequences can also be used as immunogenic compositions,
XX especially where the vector directs the expression of a neutralising
XX epitope of C. pneumoniae.
XX
XX Sequence 181 AA;
XX
XX Query Match 90.9%; Score 30; DB 20; Length 181;
XX Best Local Similarity 80.0%; Pred. No. 3.3e+02; Mismatches 0; Gaps 0;
XX Matches 4; Conservative 1; Indels 0;
XX
XX QY 1 GYXMS 5
XX |||:|
XX Db 48 GYXLS 52
XX
XX RESULT 9
XX AAU17327
XX ID AAU17327 standard; Protein; 210 AA.
XX AC AAU17327;
XX
XX 07-NOV-2001 (first entry)
XX
XX Novel signal transduction pathway protein, Seq ID 892.
XX
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
XX antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
XX immune system disorder; rheumatoid arthritis; inflammatory condition;
XX organ transplant rejection; infection; hepatitis C; blood disorder;
XX sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
XX reproductive system; gastrointestinal; liver disorder; AIDS;
XX acquired immune deficiency syndrome.
XX
XX Homo sapiens.
XX
XX WO200154733-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01312.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184684.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX

```

PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0218680.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0227009.  
 PR 01-SEP-2000; 2000US-0228287.  
 PR 01-SEP-2000; 2000US-0228287.  
 PR 01-SEP-2000; 2000US-0228343.  
 PR 01-SEP-2000; 2000US-0228344.  
 PR 01-SEP-2000; 2000US-0228345.  
 PR 05-SEP-2000; 2000US-0228509.  
 PR 05-SEP-2000; 2000US-0228513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234597.  
 PR 25-SEP-2000; 2000US-0234598.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249285.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-465460/50.  
 XX N-PSDB; AAS27244.  
 XX  
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
 XX prognosing disorders related to the proteins, including cancers, immune  
 XX disorders and neuronal disorders -  
 XX  
 XX Claim 1; SEQ ID No 892; 880pp; English.  
 XX  
 XX The invention relates to novel isolated polypeptides (I), and





CC neurological diseases. The polynucleotide can be used as a marker, in  
 CC the identification of genetic disorders, and for the design of  
 CC oligonucleotides for examining expression patterns.

XX SQ Sequence 242 AA;

Query Match 90.9%; Score 30; DB 22; Length 242;  
 Best Local Similarity 80.0%; Pred. NO. 4.4e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYVMS 5  
 |||:  
 Db 224 GYVLS 228

RESULT 12

ABB72253  
 ID ABB72253 standard; Protein; 242 AA.

XX AC ABB72253;

XX DT 04-APR-2002 (first entry)

XX DE Rat protein isolated from skin cells SEQ ID NO: 393.

XX KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
 KW developmental defect; inflammatory disease; dermatological; vulnary;  
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX OS Rattus sp.

XX PN WO200190357-A1.

XX PD 29-NOV-2001.

XX PF 24-MAY-2001; 2001WO-NZ00099.

XX PR 24-MAY-2000; 2000US-206650P.

XX PR 25-JUL-2000; 2000US-221232P.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;  
 XX WPI; 2002-122020/16.

XX PT New polynucleotides and polypeptides encoded by the polynucleotides  
 PT isolated from skin cells, useful for treating skin wounds, cancers,  
 PT growth and developmental defects, inflammatory diseases, or for  
 PT modulating immune responses -

XX PS Example 2; Page 246-247; 466pp; English.

XX CC The present invention provides the protein and coding sequences of cDNAs  
 CC isolated from human, murine and rat skin cell libraries. The sequences  
 CC can be used in the development of therapeutic agents useful in the  
 CC treatment of skin diseases, including skin wounds, cancer, growth  
 CC defects, developmental defects and inflammatory diseases. The proteins  
 CC have important roles in the induction of hair growth, cell proliferation  
 CC and cell-cell interaction, in maintaining tissue integrity, in wound  
 CC healing and in modulating immune responses. The present sequence is a  
 CC polypeptide of the invention.

XX SQ Sequence 242 AA;

Query Match 90.9%; Score 30; DB 23; Length 242;  
 Best Local Similarity 80.0%; Pred. NO. 4.4e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYVMS 5  
 |||:  
 Db 224 GYVLS 228

RESULT 13

ABB72290  
 ID ABB72290 standard; Protein; 242 AA.

XX AC ABB72290;

XX DT 04-APR-2002 (first entry)

XX DE Rat protein isolated from skin cells SEQ ID NO: 502.

XX KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
 KW developmental defect; inflammatory disease; dermatological; vulnary;  
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX OS Rattus sp.

XX PN WO200190357-A1.

XX PD 29-NOV-2001.

XX PF 24-MAY-2001; 2001WO-NZ00099.

XX PR 24-MAY-2000; 2000US-206650P.

XX PR 25-JUL-2000; 2000US-221232P.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;  
 XX WPI; 2002-122020/16.

XX PT New polynucleotides and polypeptides encoded by the polynucleotides  
 PT isolated from skin cells, useful for treating skin wounds, cancers,  
 PT growth and developmental defects, inflammatory diseases, or for  
 PT modulating immune responses -

XX PS Claim 4; Page 308; 466pp; English.

XX CC The present invention provides the protein and coding sequences of cDNAs  
 CC isolated from human, murine and rat skin cell libraries. The sequences  
 CC can be used in the development of therapeutic agents useful in the  
 CC treatment of skin diseases, including skin wounds, cancer, growth  
 CC defects, developmental defects and inflammatory diseases. The proteins  
 CC have important roles in the induction of hair growth, cell proliferation  
 CC and cell-cell interaction, in maintaining tissue integrity, in wound  
 CC healing and in modulating immune responses. The present sequence is a  
 CC polypeptide of the invention.

XX SQ Sequence 242 AA;

Query Match 90.9%; Score 30; DB 23; Length 242;  
 Best Local Similarity 80.0%; Pred. NO. 4.4e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYVMS 5  
 |||:  
 Db 224 GYVLS 228

RESULT 14

ABB72301  
 ID ABB72301 standard; Protein; 242 AA.

XX AC ABB72301;

XX DT 04-APR-2002 (first entry)

XX DE Rat protein isolated from skin cells SEQ ID NO: 625.

XX KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
 KW developmental defect; inflammatory disease; dermatological; vulnary;  
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX OS Rattus sp.  
XX PN WO200190357-A1.  
XX PD 29-NOV-2001.  
XX PF 24-MAY-2001; 2001WO-NZ00099.  
XX PR 24-MAY-2000; 2000US-206650P.  
XX PR 25-JUL-2000; 2000US-221332P.  
XX PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX PI Watson JD, Strachan L, Sleeman M, Orrust R, Murison JG, Kumble KD;  
XX DR WPI: 2002-122020/16.  
XX DR N-PSDB; ABL34986.  
XX PT New polynucleotides and polypeptides encoded by the polynucleotides  
PT isolated from skin cells, useful for treating skin wounds, cancers,  
PT growth and developmental defects, inflammatory diseases, or for  
PT modulating immune responses -  
XX PS Claim 4; Page 390; 466pp; English.  
XX SS The present invention provides the protein and coding sequences of cDNAs  
CC isolated from human, murine and rat skin cell libraries. The sequences  
CC can be used in the development of therapeutic agents useful in the  
CC treatment of skin diseases, including skin wounds, cancer, growth  
CC defects, developmental defects and inflammatory diseases. The proteins  
CC have important roles in the induction of hair growth, cell proliferation  
CC and cell-cell interaction, in maintaining tissue integrity, in wound  
CC healing and in modulating immune responses. The present sequence is a  
CC polypeptide of the invention.  
XX SQ Sequence 242 AA;  
  
Query Match 90.9%; Score 30; DB 23; Length 242;  
Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GYVMS 5  
Db 224 GYVLS 228  
  
RESULT 15  
ID ABP39269  
XX ABP39269 standard; Protein; 243 AA.  
XX AC ABP39269;  
XX DT 24-JUL-2002 (first entry)  
XX DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4114.  
XX KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
KW antibacterial; gene therapy.  
XX OS Staphylococcus epidermidis.  
XX PN US6380370-B1.  
XX PD 30-APR-2002.  
XX PF 13-AUG-1998; 98US-0134001.  
XX PR 14-AUG-1997; 97US-055779P.  
XX PR 08-NOV-1997; 97US-064964P.  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX

PI Doucette-Stamm LA, Bush D;  
XX WPI: 2002-381255/41.  
DR N-PSDB; ABN91814.  
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis  
PT polypeptide, useful for diagnosing and treating bacterial infections -  
XX Disclosure; SEQ ID 4114; 267pp; English.  
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have  
CC antibacterial activity and can be used in gene therapy. The sequences  
CC can also be used in the diagnosis and treatment of bacterial infections,  
CC particularly S. epidermidis infections. The sequences can be used to  
CC screen for compounds able to interfere with the S. epidermidis life  
CC cycle or inhibit S. epidermidis infection.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site.  
XX SQ Sequence 243 AA;  
  
Query Match 90.9%; Score 30; DB 23; Length 243;  
Best Local Similarity 80.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GYVMS 5  
Db 52 GYVLS 56  
  
Search completed: April 22, 2003, 12:51:15  
Job time : 14.6786 secs



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	30	90.9	64	9	US-10-144-929-162	Sequence 162, App
2	30	90.9	210	9	US-09-764-868-892	Sequence 892, App
3	30	90.9	242	9	US-09-866-050A-393	Sequence 393, App
4	30	90.9	242	9	US-09-866-050A-502	Sequence 502, App
5	30	90.9	242	9	US-09-866-050A-625	Sequence 625, App
6	30	90.9	310	10	US-09-815-242-10910	Sequence 10910, App
7	30	90.9	392	9	US-10-050-704-180	Sequence 180, App
8	30	90.9	392	9	US-10-144-929-156	Sequence 156, App
9	30	90.9	420	9	US-09-905-231A-109	Sequence 109, App
10	30	90.9	420	9	US-09-902-853-109	Sequence 109, App
11	30	90.9	420	9	US-09-907-824-109	Sequence 109, App
12	30	90.9	420	9	US-09-907-841-109	Sequence 109, App
13	30	90.9	420	9	US-09-904-011-109	Sequence 109, App
14	30	90.9	420	9	US-09-908-011-109	Sequence 109, App
15	30	90.9	420	9	US-10-028-072-290	Sequence 290, App
16	30	90.9	420	9	US-09-906-742-109	Sequence 290, App
17	30	90.9	420	9	US-10-121-049-290	Sequence 290, App
18	30	90.9	420	9	US-10-123-094-290	Sequence 290, App
19	30	90.9	420	9	US-10-140-470-290	Sequence 290, App
20	30	90.9	420	9	US-09-796-753-84	Sequence 84, App1

RESULT 2  
US-03-764-868-892  
; Sequence 892, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```

; FILE REFERENCE: PT332
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 892
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-892

```

```

Query Match          90.9%; Score 10; DB 9; Length 210;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GYWMS 5
   |||:|
Db 68 GYWLS 72

```

```

RESULT 3
US-09-866-050A-393
; Sequence 393, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866.050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 393
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-393

```

```

Query Match          90.9%; Score 30; DB 9; Length 242;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GYWMS 5
   |||:|
Db 224 GYWLS 228

```

```

RESULT 4
US-09-866-050A-502
; Sequence 502, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866.050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 502

```

```

; LENGTH: 242
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-502

```

```

Query Match          90.9%; Score 30; DB 9; Length 242;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GYWMS 5
   |||:|
Db 224 GYWLS 228

```

```

RESULT 5
US-09-866-050A-625
; Sequence 625, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866.050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 625
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-625

```

```

Query Match          90.9%; Score 30; DB 9; Length 242;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GYWMS 5
   |||:|
Db 224 GYWLS 228

```

```

RESULT 6
US-09-815-242-10910
; Sequence 10910, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23

```

Thu Apr 24 09:14:10 2003

us-09-674-716b-9.open.rapb

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10910
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10910
```

```
Query Match 90.9%; Score 30; DB 10; Length 310;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GYWMS 5
    |||:|
Db 87 GYWLS 91
```

```
RESULT 7
US-10-050-704-180
; Sequence 180, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 180
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (251)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-050-704-180
```

```
Query Match 90.9%; Score 30; DB 9; Length 392;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GYWMS 5
    |||:|
Db 374 GYWLS 378
```

```
RESULT 8
US-10-144-929-156
; Sequence 156, Application US/10144929
; Publication No. US20030069405A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2014P1
; CURRENT APPLICATION NUMBER: US/10/144,929
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/251,329
```

```
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: PCT/US98/17044
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 156
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (251)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-144-929-156
```

```
Query Match 90.9%; Score 30; DB 9; Length 392;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GYWMS 5
    |||:|
Db 374 GYWLS 378
```

```
RESULT 9
US-09-905-291A-109
; Sequence 109, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 109
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-109

Query Match          90.9%; Score 30; DB 9; Length 420;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWS 5
    |||||
Db 402 GYWS 406

RESULT 10
US-09-902-853-109
; Sequence 109, Application US/09902853
; Publication No. US20020192859A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18

```

```

; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 109
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-853-109

Query Match          90.9%; Score 30; DB 9; Length 420;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWS 5
    |||||
Db 402 GYWS 406

RESULT 11
US-09-907-824-109
; Sequence 109, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.

```



```

; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 109
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-109

```

```

Query Match          90.9%; Score 30; DB 9; Length 420;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GYWS 5
      |||||
Db      402 GYWS 406

```

```

RESULT 12
US-09-907-841-109
; Sequence 109, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen

```

```

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 109
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-841-109

```

```

Query Match          90.9%; Score 30; DB 9; Length 420;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GYWS 5
      |||||
Db      402 GYWS 406

```

```

RESULT 13
US-09-904-011-109
; Sequence 109, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen

```

```

; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gottard, A.
; APPLICANT: Goddard, A.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Hillan, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 109
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-109

Query Match 90.9%; Score 30; DB 9; Length 420;
Best Local Similarity 80.0%; Pred No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5
Db 402 GYWS 406

```

```

RESULT 14
US-10-028-072-290
; Sequence 290, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561

```

;; PRIOR FILING DATE: 1997-10-28  
;; PRIOR APPLICATION NUMBER: 60/063704  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063733  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063735  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063738  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063755  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/064248  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/064809  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/065186  
;; PRIOR FILING DATE: 1997-11-12  
;; PRIOR APPLICATION NUMBER: 60/065846  
;; PRIOR FILING DATE: 1997-11-17  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/066453  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066511  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/069212  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069278  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069334  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/072320  
;; PRIOR FILING DATE: 1998-01-23  
;; PRIOR APPLICATION NUMBER: 60/073612  
;; PRIOR FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-02-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081695  
;; PRIOR FILING DATE: 1998-04-14  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07

;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07

Query Match 90.9%; Score 30; DB 9; Length 420;  
Best Local Similarity 80.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||  
Db 402.GYWLS 406

RESULT 15  
US-09-906-742-109  
; Sequence 109, Application US/09906742  
; Publication No. US20030023054A1  
; GENERAL INFORMATION:

APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherwan  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/906,742  
PRIOR FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 09/665,350  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 109  
LENGTH: 420  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-906-742-109

Query Match 90.9%; Score 30; DB 9; Length 420;  
Best Local Similarity 80.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5  
|||:  
Db 402 GYWLS 406

Search completed: April 22, 2003, 13:11:14  
Job time : 6.26786 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:49 ; Search time 4.375 Seconds  
(without alignments)  
33.626 Million cell updates/sec

Title: US-09-674-716B-9  
Perfect score: 33  
Sequence: 1 GYMS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PCRU COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	90.9	43	2	US-08-488-161-39
2	30	90.9	43	3	US-09-273-685-39
3	30	90.9	43	5	PCT-US95-11934-39
4	30	90.9	243	4	US-09-134-001C-4114
5	30	90.9	667	4	US-09-315-127-5
6	30	90.9	667	4	US-09-315-127-6
7	29	87.9	5	2	US-08-480-434-22
8	29	87.9	5	2	US-08-480-434-31
9	29	87.9	5	2	US-08-053-451B-22
10	29	87.9	5	2	US-08-053-451B-31
11	29	87.9	14	1	US-08-204-656B-11
12	29	87.9	14	1	US-08-470-702-11
13	29	87.9	14	1	US-08-467-831-11
14	29	87.9	20	2	US-08-053-451B-114
15	29	87.9	119	2	US-08-053-451B-125
16	29	87.9	121	2	US-08-480-434-29
17	29	87.9	121	2	US-08-053-451B-29
18	29	87.9	126	2	US-08-480-434-18
19	29	87.9	126	2	US-08-480-434-19
20	29	87.9	126	2	US-08-053-451B-18
21	29	87.9	126	2	US-08-053-451B-19
22	29	87.9	679	3	US-09-257-536-2
23	29	87.9	679	4	US-09-512-230-2
24	28	84.8	274	4	US-09-570-367C-19
25	28	84.8	482	4	US-09-134-001C-4309
26	27	81.8	5	1	US-08-244-626-12
27	27	81.8	5	2	US-08-480-434-32

28	27	81.8	5	2	US-08-318-157B-23	Sequence 23, Appl
29	27	81.8	5	2	US-08-053-451B-32	Sequence 32, Appl
30	27	81.8	24	4	US-08-525-539A-19	Sequence 19, Appl
31	27	81.8	98	1	US-07-942-245-37	Sequence 37, Appl
32	27	81.8	99	1	US-08-244-626-10	Sequence 10, Appl
33	27	81.8	109	1	US-07-942-245-17	Sequence 17, Appl
34	27	81.8	115	4	US-09-376-330-20	Sequence 20, Appl
35	27	81.8	116	3	US-08-483-749A-18	Sequence 18, Appl
36	27	81.8	117	1	US-07-958-140-3	Sequence 3, Appl
37	27	81.8	117	3	US-08-545-809A-95	Sequence 95, Appl
38	27	81.8	117	5	PCT-US93-09166-3	Sequence 3, Appl
39	27	81.8	119	2	US-08-318-157B-2	Sequence 2, Appl
40	27	81.8	119	2	US-08-318-157B-8	Sequence 8, Appl
41	27	81.8	119	2	US-08-318-157B-9	Sequence 9, Appl
42	27	81.8	119	2	US-08-318-157B-10	Sequence 10, Appl
43	27	81.8	119	2	US-08-318-157B-11	Sequence 11, Appl
44	27	81.8	119	2	US-08-318-157B-12	Sequence 12, Appl
45	27	81.8	119	2	US-08-318-157B-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-08-488-161-39  
; Sequence 39, Application US/08488161  
; Patent No. 5885577  
; \*GENERAL INFORMATION:  
; APPLICANT: Alvarez, Vernon L.  
; TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,161  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Marock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: amino acid  
; LENGTH: 43 amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-488-161-39

Query Match 90.9%; Score 30; DB 2; Length 43;  
Best Local Similarity 80.0%; Pred. No. 45;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GYMS 5  
DB 6 GYMLS 10

```
RESULT 2
US-09-273-685-39
; Sequence 39, Application US/09273685
; Patent No. 6015561
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Vernon L.
; TITLE OF INVENTION: Antigen Binding Peptides (Aptides) From
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; PRIOR APPLICATION NUMBER: US/09/273,685
; CLASSIFICATION:
; APPLICATION NUMBER: 08/488,161
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-176
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-273-685-39
Query Match 90.9%; Score 30; DB 3; Length 43;
Best Local Similarity 80.0%; Pred. No. 45;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWMS 5
Db 6 GYWLS 10
RESULT 3
PCT-US95-11934-39
; Sequence 39, Application PC/TUS9511934
; GENERAL INFORMATION:
; APPLICANT: Cycogen Corporation
; TITLE OF INVENTION: Antigen Binding Peptides (Aptides) From
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11934
; FILING DATE: 20-SEP-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-196-228
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-11934-39
Query Match 90.9%; Score 30; DB 5; Length 43;
Best Local Similarity 80.0%; Pred. No. 45;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWMS 5
Db 6 GYWLS 10
RESULT 4
US-09-134-001C-4114
; Sequence 4114, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lyda Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4114
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4114
Query Match 90.9%; Score 30; DB 4; Length 243;
Best Local Similarity 80.0%; Pred. No. 2,3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWMS 5
Db 52 GYWLS 56
RESULT 5
US-09-315-127-5
; Sequence 5, Application US/09315127
; Patent No. 6448390
; GENERAL INFORMATION:
; APPLICANT: The University of Tennessee, c/o Richard Cox
; TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and
; TITLE OF INVENTION: Liposome Vectors and Use in Gene and Drug Therapy
; FILE REFERENCE: 44137-5023, U. of Tennessee
; CURRENT APPLICATION NUMBER: US/09/315,127
```

; CURRENT FILING DATE: 1999-05-20  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 667  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 US-09-315-127-5

Query Match 90.9%; Score 30; DB 4; Length 667;  
 Best Local Similarity 80.0%; Pred. No. 5.8e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWS 5  
 Db 180 GYWS 184

# RESULT 6

US-09-315-127-6  
 ; Sequence 6, Application US/09315127  
 ; Patent No. 6448390  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The University of Tennessee, c/o Richard Cox  
 ; TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and  
 ; FILE REFERENCE: 44137-5023, U. of Tennessee  
 ; CURRENT APPLICATION NUMBER: US/09/315,127  
 ; CURRENT FILING DATE: 1999-05-20  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 6  
 ; LENGTH: 667  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: SEQ ID NO. 4,  
 ; OTHER INFORMATION: envelope protein produced by retroviral vector of  
 ; OTHER INFORMATION: seq. id no. 3  
 US-09-315-127-6

Query Match 90.9%; Score 30; DB 4; Length 667;  
 Best Local Similarity 80.0%; Pred. No. 5.8e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWS 5  
 Db 180 GYWS 184

# RESULT 7

US-08-480-434-22  
 ; Sequence 22, Application US/08480434  
 ; Patent No. 5811248  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles C. Ditlow, et al.  
 ; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC ANTIGENS,  
 ; TITLE OF INVENTION: ANTIBODIES THERETO, AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 88  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.24  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/480,434

; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Albert P. Halluin  
 ; REGISTRATION NUMBER: 25,227  
 ; REFERENCE/DOCKET NUMBER: 7606-053  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 854-3660  
 ; TELEFAX: (415) 854-3694  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 22:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: N  
 ; ANTI-SENSE: N  
 US-08-480-434-22

Query Match 87.9%; Score 29; DB 2; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 2e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWS 5  
 Db 1 GYWS 5

# RESULT 8

US-08-480-434-31  
 ; Sequence 31, Application US/08480434  
 ; Patent No. 5811248  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles C. Ditlow, et al.  
 ; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC ANTIGENS,  
 ; TITLE OF INVENTION: ANTIBODIES THERETO, AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 88  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.24  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/480,434  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Albert P. Halluin  
 ; REGISTRATION NUMBER: 25,227  
 ; REFERENCE/DOCKET NUMBER: 7606-053  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 854-3660  
 ; TELEFAX: (415) 854-3694  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 31:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: N  
 ; ANTI-SENSE: N  
 US-08-480-434-31

```

Query Match      87.9%; Score 29; DB 2; Length 5;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5
Db 1 GFWS 5

RESULT 9
US-08-053-451B-22
; Sequence 22, Application US/08053451B
; Patent No. 5955584
; GENERAL INFORMATION:
; APPLICANT: Chen, Francis W.
; APPLICANT: Dittlow, Charles C.
; APPLICANT: Calenoff, Emanuel
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/053.451B
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 7606-033-999
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3660
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; HYPOTHETICAL: N
; ANTI-SENSE: N
; US-08-053-451B-22

Query Match      87.9%; Score 29; DB 2; Length 5;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5
Db 1 GFWS 5

RESULT 10
US-08-053-451B-31
; Sequence 31, Application US/08053451B
; Patent No. 5955584
; GENERAL INFORMATION:
; APPLICANT: Chen, Francis W.
; APPLICANT: Dittlow, Charles C.
; APPLICANT: Calenoff, Emanuel

```

```

; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC
; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/053.451B
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 7606-033-999
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3660
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; HYPOTHETICAL: N
; ANTI-SENSE: N
; US-08-053-451B-31

Query Match      87.9%; Score 29; DB 2; Length 5;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWMS 5
Db 1 GFWS 5

RESULT 11
US-08-204-656B-11
; Sequence 11, Application US/08204656B
; Patent No. 5538882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuo
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```



APPLICATION NUMBER: US/08/204,656B  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WEINER, MARC S.  
REGISTRATION NUMBER: 32,181  
REFERENCE/DOCKET NUMBER: 234-252P  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-204-656B-11

Query Match 87.9%; Score 29; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWM 4  
|||||  
DB 10 GYWM 13

## RESULT 12

US-08-470-702-11  
Sequence 11, Application US/08470702  
Patent No. 5631149  
GENERAL INFORMATION:  
APPLICANT: MATSUI, IKUO  
APPLICANT: ISHIKAWA, KAZUHIKO  
APPLICANT: MIYAIRI, SACHIO  
APPLICANT: HONDA, KOICHI  
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,  
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING  
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 8110 Gatehouse Road, Suite 500 East  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22042  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,702  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/204,656  
FILING DATE: 02-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: WEINER, MARC S.  
REGISTRATION NUMBER: 32,181  
REFERENCE/DOCKET NUMBER: 234-252P  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-470-702-11

Query Match 87.9%; Score 29; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWM 4  
|||||  
DB 10 GYWM 13

## RESULT 13

US-08-467-831-11  
Sequence 11, Application US/08467831  
Patent No. 5635378  
GENERAL INFORMATION:  
APPLICANT: MATSUI, IKUO  
APPLICANT: ISHIKAWA, KAZUHIKO  
APPLICANT: MIYAIRI, SACHIO  
APPLICANT: HONDA, KOICHI  
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,  
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING  
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 8110 Gatehouse Road, Suite 500 East  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22042  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,831  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/204,656  
FILING DATE: 02-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: WEINER, MARC S.  
REGISTRATION NUMBER: 32,181  
REFERENCE/DOCKET NUMBER: 234-252P  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-467-831-11

Query Match

87.9%; Score 29; DB 1; Length 14;

Best Local Similarity 100.0%; Pred. No. 23;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWM 4  
|:|:|  
Db 10 GYWM 13

## RESULT 14

US-08-053-451B-114  
; Sequence 114, Application US/08053451B  
; Patent No. 5955584  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Francis W.  
; APPLICANT: Ditlow, Charles C.  
; APPLICANT: Calenoff, Emanuel  
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC  
; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF  
; NUMBER OF SEQUENCES: 176  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/053,451B  
; FILING DATE: 26-APR-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halluin, Albert P.  
; REGISTRATION NUMBER: 25,227  
; REFERENCE/DOCKET NUMBER: 7606-033-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-854-3660  
; TELEFAX: 415-854-3694  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-053-451B-114

Query Match 87.9%; Score 29; DB 2; Length 20;  
Best Local Similarity 80.0%; Pred. No. 33;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWM 5  
|:|:|  
Db 10 GYWM 14

## RESULT 15

US-08-053-451B-125  
; Sequence 125, Application US/08053451B  
; Patent No. 5955584  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Francis W.  
; APPLICANT: Ditlow, Charles C.  
; APPLICANT: Calenoff, Emanuel  
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC  
; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF  
; NUMBER OF SEQUENCES: 176  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/053,451B  
; FILING DATE: 26-APR-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halluin, Albert P.  
; REGISTRATION NUMBER: 25,227  
; REFERENCE/DOCKET NUMBER: 7606-033-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-854-3660  
; TELEFAX: 415-854-3694  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 125:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 119 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-053-451B-125

Query Match 87.9%; Score 29; DB 2; Length 119;  
Best Local Similarity 80.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWM 5  
|:|:|  
Db 31 GYWM 35

Search completed: April 22, 2003, 12:55:42  
Job time : 5.375 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:24 ; Search time 18.6607 Seconds  
(without alignments)  
97.882 Million cell updates/sec

Title: US-09-674-716B-11  
 Perfect score: 98  
 Sequence: 1 BIRLSDNYATHAESVKG 19  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 2832224

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : PIR 73:*
1: _pir1:
2: _pir2:
3: _pir3:
4: _pir4:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	98	100.0	101	2	S26460	Ig heavy chain V r	
2	98	100.0	115	2	S38714	Ig heavy chain V r	
3	93	94.9	106	2	S24521	Ig heavy chain V r	
4	93	94.9	115	1	AVMS06	Ig heavy chain V-I	
5	91	92.9	113	1	AVMSAB	Ig heavy chain V-I	
6	91	92.9	113	1	AVMS61	Ig heavy chain V-I	
7	91	92.9	113	1	AVMS09	Ig heavy chain V-I	
8	91	92.9	113	1	AVMSB7	Ig heavy chain V-I	
9	91	92.9	115	1	AVMS82	Ig heavy chain V-I	
10	89	90.8	115	2	A25803	Ig heavy chain V r	
11	86	87.8	113	1	AVMS57	Ig heavy chain V r	
12	85	86.7	113	1	HVMSAM	Ig heavy chain V-I	
13	75.5	77.0	111	1	HVMS76	Ig heavy chain V-I	
14	73	74.5	139	2	PCJ213	Ig heavy chain pre	
15	68	69.4	64	2	I77394	Ig heavy chain V r	
16	68	69.4	121	2	S09958	Ig heavy chain V-D	
17	67	68.4	100	2	S26462	Ig heavy chain V r	
18	67	68.4	110	2	PHI091	Ig heavy chain V r	
19	67	68.4	110	2	PHI092	Ig heavy chain V r	
20	67	68.4	119	2	A27630	Ig heavy chain pre	
21	67	68.4	121	2	A41940	Ig heavy chain V r	
22	67	68.4	141	2	I32513	Ig heavy chain pre	
23	66	67.3	135	2	S67945	Ig heavy chain BrE	
24	65	66.3	139	2	G29380	Ig heavy chain pre	
25	64	65.3	110	2	PHI093	Ig heavy chain V r	
26	63	64.3	220	2	S68211	Ig heavy chain (Ma	
27	63	64.3	444	2	PC4436	monoclonal antibody	
28	62	62.2	119	2	PHI293	Ig heavy chain pre	
29	61	62.2	137	2	S42467	Ig heavy chain V r	

## ALIGNMENTS

## RESULT 1

S26460  
Ig heavy chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S26460

R;Kavaler, J.  
submitted to the EMBL Data Library April 1991

A:Reference number: S26459

A;Reference number: S26460  
A;Accession: S26460

A;Status: preliminary

A; Molecule type: mRNA

A;Residues: 1-101 &lt;KAV&gt;

A; Cross-references: EMBL:X559106; NID:951707; PDB:CAA41832.1; FID:931708

C; Superfamily: immunoglobulin V region; Immunoglobulin  
C; Superfamily: heterotetramer: immunoglobulin  
C; Keywords:

**C;keywords:** heterocellulamer; immunoglobulin

Query Match 100.0%; Score 98; DB 2; Length 101;

Best Local Similarity 100.0%; Pred. No. 1.1e-08;

Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

Qy 1 EIRKSDNYATHAESVKG 19

100

## RESULT 2

S38714  
Ig heavy chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Jun-2000  
C:Accession: S38714  
R:Cimanis, A.Y.

R;Clmanis, A.Y.  
submitted to the EMBL Data Library, November 1993

A:Reference number: S38713

A;Accession: S38714

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-115 <C1M>

A; Cross-references: EMBL:X/6014; NID:G416092; FID:CA33601.1; FID:G416092.1; C.superfamily: immunoglobulin V region: immunoglobulin homology

**C;superfamily:** immunoglobulin V region; immunoglobulin C;  
**C;keywords:** heterotetramer; immunoglobulin

F:14-99/Domain: immunoglobulin homology <IMM>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Query Match 100.0%; Score 98; DB 2; Length 115;

Matches 19; Conservative

Q: 1 E1P1 KODNNVATUNAFSIVK 19

## RESULT 3

S24521  
Ig heavy chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999  
C:Accession: S24521  
R:Kaartinen, M.  
submitted to the EMBL Data Library, October 1991  
A:Reference number: S24490  
A:Accession: S24521  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-106 <KAA>  
C:Cross-references: EMBL:X66664; NID:G51247; PID:CAA47226.1; PID:G51248  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:15-106/Domain: immunoglobulin homology <IMM>

Query Match 94.9%; Score 93; DB 2; Length 106;  
Best Local Similarity 94.7%; Pred. No. 7.5e-08;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRLKSNDNYATHYAESVKG 19  
|||||:|||||

Db 56 EIRLKSNNYATHYAESVKG 74

## RESULT 4

AVMS06  
Ig heavy chain V-III region (J606) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Aug-1996  
C:Accession: C92811; A02072  
R:Johnson, N.; Sliankard, J.; Paul, L.; Hood, L.  
J. Immunol. 128, 302-307, 1982  
A:Title: The complete V domain amino acid sequences of two myeloma inulin-binding proteins.  
A:Reference number: A92811; MUID:82099361; PMID:6798111  
A:Accession: C92811  
A:Molecule type: protein  
A:Residues: 1-115 <JOH>  
C:Comment: This chain was isolated from a myeloma protein that binds inulin.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:15-100/Domain: immunoglobulin homology <IMM>  
F:22-98/Disulfide bonds: #status predicted

Query Match 94.9%; Score 93; DB 1; Length 115;  
Best Local Similarity 94.7%; Pred. No. 8.2e-08;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRLKSNDNYATHYAESVKG 19  
|||||:|||||

Db 50 EIRLKSNNYATHYAESVKG 68

## RESULT 5

AVNSAE  
Ig heavy chain V-III region (A4) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 24-Apr-1984 #sequence\_revision 30-Jun-1993 #text\_change 31-Mar-1997  
C:Accession: A93818; A02072  
R:Vrana, M.; Rudikoff, S.; Potter, M.  
Proc. Natl. Acad. Sci. U.S.A. 75, 1957-1961, 1978  
A:Title: Sequence variation among heavy chains from inulin-binding myeloma proteins.  
A:Reference number: A93818; MUID:78158406; PMID:417344  
A:Accession: A93818  
A:Molecule type: protein  
A:Residues: 1-113 <VRA>  
C:Comment: This chain was isolated from a myeloma protein that binds inulin.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:15-100/Domain: immunoglobulin homology <IMM>

Query Match 92.9%; Score 91; DB 1; Length 113;  
Best Local Similarity 94.7%; Pred. No. 1.7e-07;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSNDNYATHYAESVKG 19  
|||||:|||||

Db 50 EIRLKSNNYATHYAESVKG 68

## RESULT 6

AVMS61  
Ig heavy chain V-III region (U61) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Aug-1996  
C:Accession: B93818; A02072  
R:Vrana, M.; Rudikoff, S.; Potter, M.  
Proc. Natl. Acad. Sci. U.S.A. 75, 1957-1961, 1978  
A:Title: Sequence variation among heavy chains from inulin-binding myeloma proteins.  
A:Reference number: A93818; MUID:78158406; PMID:417344  
A:Accession: B93818  
A:Molecule type: protein  
A:Residues: 1-113 <VRA>  
C:Comment: This chain was isolated from a myeloma protein that binds inulin.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:15-100/Domain: immunoglobulin homology <IMM>  
F:22-98/Disulfide bonds: #status predicted

F:22-98/Disulfide bonds: #status predicted

Query Match 92.9%; Score 91; DB 1; Length 113;  
Best Local Similarity 94.7%; Pred. No. 1.7e-07;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSNDNYATHYAESVKG 19  
|||||:|||||

Db 50 EIRLKSNNYATHYAESVKG 68

## RESULT 7

AVMS09  
Ig heavy chain V-III region (E109) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Aug-1996  
C:Accession: C93818; A02072  
R:Vrana, M.; Rudikoff, S.; Potter, M.  
Proc. Natl. Acad. Sci. U.S.A. 75, 1957-1961, 1978  
A:Title: Sequence variation among heavy chains from inulin-binding myeloma proteins.  
A:Reference number: A93818; MUID:78158406; PMID:417344  
A:Accession: C93818  
A:Molecule type: protein  
A:Residues: 1-113 <VRA>  
C:Comment: This chain was isolated from a myeloma protein that binds inulin.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:15-100/Domain: immunoglobulin homology <IMM>  
F:22-98/Disulfide bonds: #status predicted

Query Match 92.9%; Score 91; DB 1; Length 113;  
Best Local Similarity 94.7%; Pred. No. 1.7e-07;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSNDNYATHYAESVKG 19  
|||||:|||||

Db 50 EIRLKSNNYATHYAESVKG 68

## RESULT 8

AVMSB7  
Ig heavy chain V-III region (ABE-47N) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Aug-1996  
C:Accession: A90400; A02072  
R:Vrana, M.; Rudikoff, S.; Potter, M.  
Proc. Natl. Acad. Sci. U.S.A. 75, 1957-1961, 1978  
A:Title: Sequence variation among heavy chains from inulin-binding myeloma proteins.  
A:Reference number: A93818; MUID:78158406; PMID:417344  
A:Accession: C93818  
A:Molecule type: protein  
A:Residues: 1-113 <VRA>  
C:Comment: This chain was isolated from a myeloma protein that binds inulin.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:15-100/Domain: immunoglobulin homology <IMM>  
F:22-98/Disulfide bonds: #status predicted

Query Match 92.9%; Score 91; DB 1; Length 113;  
Best Local Similarity 94.7%; Pred. No. 1.7e-07;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSNDNYATHYAESVKG 19  
|||||:|||||

Db 50 EIRLKSNNYATHYAESVKG 68

## RESULT 9

AVMSB7  
Ig heavy chain V-III region (ABE-47N) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Aug-1996  
C:Accession: A90400; A02072

R;Vrana, M.; Rudikoff, S.; Potter, M.

Biochemistry 16, 1170-1175, 1977

A;Title: Heavy-chain variable-region sequence from an inulin-binding myeloma protein.

A;Reference number: A90400; MUID:77134726; PMID:402936

A;Accession: A90400

A;Molecule type: protein

A;Residues: 1-113 <VRA>

C;Comment: This chain was isolated from a myeloma protein that binds inulin.

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-100/Domain: immunoglobulin homology <IMM>

F;22-98/Disulfide bonds: #status predicted

Query Match 92.9%; Score 91; DB 1; Length 113;

Best Local Similarity 94.7%; Pred. No. 1.7e-07;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19

||||| |||||||

Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 9

AVMS82

Ig heavy chain V-III region (W3082) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Aug-1996

C;Accession: D92811; A02072

R;Johnson, N.; Slankard, J.; Paul, L.; Hood, L.

J. Immunol. 128, 302-307, 1982

A;Title: The complete V domain amino acid sequences of two myeloma inulin-binding proteins

A;Reference number: A92811; MUID:82099361; PMID:679811

A;Accession: D92811

A;Molecule type: protein

A;Residues: 1-115 <JOH>

C;Comment: This chain was isolated from a myeloma protein that binds inulin.

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-100/Domain: immunoglobulin homology <IMM>

F;22-98/Disulfide bonds: #status predicted

Query Match 92.9%; Score 91; DB 1; Length 115;

Best Local Similarity 94.7%; Pred. No. 1.7e-07;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19

||||| |||||||

Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 10

A25803

Ig heavy chain V region (281.3) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 16-Aug-1996

C;Accession: A25803

R;Herbat, H.; Grutter, T.; Aebersold, R.; Braun, D.G.

Biol. Chem. Hoppe-Seyler 367, 843-851, 1986

A;Title: The complete amino-acid sequence of the variable domain of a monoclonal anti-sc

A;Reference number: A25803; MUID:87076047; PMID:3539142

A;Accession: A25803

A;Molecule type: protein

A;Residues: 1-115 <HER>

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-100/Domain: immunoglobulin homology <IMM>

Query Match 90.8%; Score 89; DB 2; Length 115;

Best Local Similarity 89.5%; Pred. No. 3.6e-07;

Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19

:||| |||||||

Db 50 QIRLASDNYATHYAESVKG 68

RESULT 11

AVMS57

Ig heavy chain V-III region (T957) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Aug-1996

C;Accession: A92810; A02072

R;Rudikoff, S.; Potter, M.

J. Immunol. 127, 191-194, 1981

A;Title: Immunoglobulin heavy chains from anti-inulin myeloma proteins: evidence for a n

A;Reference number: A92810; MUID:81216632; PMID:6787122

A;Accession: A92810

A;Molecule type: protein

A;Residues: 1-113 <RUD>

C;Comment: This chain was isolated from a myeloma protein that binds inulin.

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-100/Domain: immunoglobulin homology <IMM>

F;22-98/Disulfide bonds: #status predicted

Query Match 87.8%; Score 86; DB 1; Length 113;

Best Local Similarity 89.5%; Pred. No. 1.1e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19

||||| |||||||

Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 12

HVMSAM

Ig heavy chain V region (AMPC1) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 31-Mar-1997

C;Accession: A02073

R;Rudikoff, S.; Potter, M.

J. Immunol. 127, 191-194, 1981

A;Title: Immunoglobulin heavy chains from anti-inulin myeloma proteins: evidence for a n

A;Reference number: A92810; MUID:81216632; PMID:6787122

A;Accession: A02073

A;Molecule type: protein

A;Residues: 1-113 <RUD>

C;Comment: This chain was isolated from a myeloma protein that binds inulin.

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-100/Domain: immunoglobulin homology <IMM>

F;22-98/Disulfide bonds: #status predicted

Query Match 86.7%; Score 85; DB 1; Length 113;

Best Local Similarity 89.5%; Pred. No. 1.6e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19

||||| |||||||

Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 13

MMMS76

Ig heavy chain V-III region (HPC76) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1980 #sequence\_revision 30-Sep-1980 #text\_change 31-Mar-1997

C;Accession: A02074

R;Bernard, O.; Gough, N.M.

Proc. Natl. Acad. Sci. U.S.A. 77, 3630-3634, 1980

A;Title: Nucleotide sequence of immunoglobulin heavy chain joining segments between tran

A;Reference number: A02074; MUID:81013937; PMID:6251474

A;Accession: A02074

A;Molecule type: mRNA

A;Residues: 1-111 <BER>

A;Note: the sequence was also determined from the differentiated gene

A>Note: the sequence of the first 197 residues of the C region was also determined and d  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:10-94/Domain: immunoglobulin homology <IMM>  
F:17-92/Disulfide bonds: #status predicted

Query Match 77.0%; Score 75.5; DB 1; Length 111;  
Best Local Similarity 89.5%; Pred. No. 5.3e-05;  
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 EIRLSDNYATHYAESVKG 19  
DB 45 EIRLKS-GYATHYAESVKG 62

## RESULT 14

PC1213  
Ig heavy chain precursor V region (mAb H8)/- mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 23-Jul-1999  
C:Accession: PC1213  
R:Hong, H.J.; Kim, A.K.; Ryu, C.J.; Park, S.S.; Chung, H.K.; Kwon, K.S.; Kim, K.L.; Kim,  
Gene 121, 331-335, 1992  
A:Title: Cloning and characterization of cDNAs coding for heavy and light chains of a m  
A:Reference number: PC1213; MUID:93077049; PMID:1446832  
A:Accession: PC1213  
A:Molecule type: mRNA  
A:Residues: 1-139 <HON>  
A:Cross-references: GB:M98041; NID:g195092; PIDN:AAA38159.1; PID:g195093  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-139/Product: Ig heavy chain V region #status predicted <MAT>  
F:34-119/Domain: immunoglobulin homology <IMM>

Query Match 74.5%; Score 73; DB 2; Length 139;  
Best Local Similarity 73.7%; Pred. NO. 0.00017;  
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19  
DB 69 EIRKANNHATHYAESVKG 87

## RESULT 15

I77394  
Ig heavy chain V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
C:Accession: I77394  
R:Owens, J.D.  
Mol. Cell. Biol. 11, 5660-5670, 1991  
A:Title: Non-homologous recombination/deletion at sites within the mouse JH-cdelta locus  
A:Reference number: I57520; MUID:92017847; PMID:1922069  
A:Accession: I77394  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-64 <RES>  
A:Cross-references: GB:M64568; NID:g198472; PIDN:AAA39341.1; PID:g198473  
C:Superfamily: immunoglobulin V region; immunoglobulin homology

Query Match 69.4%; Score 68; DB 2; Length 64;  
Best Local Similarity 72.2%; Pred. NO. 0.00047;  
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 IRLKSDNYATHYAESVKG 19  
DB 9 IIVKSDNYGYATHYAESVKG 26

Search completed: April 22, 2003, 12:54:51  
Job time : 18.6607 secs



```

RA Vrana M., Rudikoff S., Potter M.;
RT "Sequence variation among heavy chains from inulin-binding myeloma
RL proteins.";
RC Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961 (1978).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; A93818; AVMSAB.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
KW DISULFID 22 98
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12675 MW; 76658C121C598285 CRC64;

Query Match 92.9%; Score 91; DB 1; Length 113;
Best Local Similarity 94.7%; Pred. No. 2.9e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSNDYATHYAESVKG 19
Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 3
HV28_MOUSE
ID HV28_MOUSE STANDARD; PRT; 113 AA.
AC P01797;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region U61.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=78158406; PubMed=417344;
RA Vrana M., Rudikoff S., Potter M.;
RT "Sequence variation among heavy chains from inulin-binding myeloma
RL proteins.";
RC Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961 (1978).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; B93818; AVMS61.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
KW DISULFID 22 98
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12671 MW; 76634C121C598285 CRC64;

Query Match 92.9%; Score 91; DB 1; Length 113;
Best Local Similarity 94.7%; Pred. No. 2.9e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSNDYATHYAESVKG 19
Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 4
HV29_MOUSE
ID HV29_MOUSE STANDARD; PRT; 113 AA.
AC P01798;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)

```

```

DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region E109.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=78158406; PubMed=417344;
RA Vrana M., Rudikoff S., Potter M.;
RT "Sequence variation among heavy chains from inulin-binding myeloma
RL proteins.";
RC Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961 (1978).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; C93818; AVMS09.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
KW DISULFID 22 98
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12647 MW; EE50F2F20EDB129B CRC64;

Query Match 92.9%; Score 91; DB 1; Length 113;
Best Local Similarity 94.7%; Pred. No. 2.9e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSNDYATHYAESVKG 19
Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 5
HV30_MOUSE
ID HV30_MOUSE STANDARD; PRT; 113 AA.
AC P01799;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region ABE-47N.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=77134726; PubMed=402936;
RA Vrana M., Rudikoff S., Potter M.;
RT "Heavy-chain variable-region sequence from an inulin-binding myeloma
RL protein.";
RC Biochemistry 16:1170-1175 (1977).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; A90400; AVMSB7.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
KW DISULFID 22 98
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12675 MW; 76658C16C779845E CRC64;

Query Match 92.9%; Score 91; DB 1; Length 113;
Best Local Similarity 94.7%; Pred. No. 2.9e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSNDYATHYAESVKG 19

```



```

Db 50 EIRLKSHNYATHYAESVKG 68
RESULT 6
HV33 MOUSE
ID HV33 MOUSE STANDARD; PRT; 115 AA.
AC P01800;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region W3082.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RN
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Sliankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma inulin-
binding proteins."
RL J. Immunol. 128:302-307(1982).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; D92811; AVMS82.
DR HSSP; P01810; 2FBU.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12887 MW; 9B4517648C121C5A CRC64;

Query Match 92.9%; Score 91; DB 1; Length 115;
Best Local Similarity 94.7%; Pred. No. 2.9e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSDNYATHYAESVKG 19
Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 7
HV31 MOUSE
ID HV31 MOUSE STANDARD; PRT; 113 AA.
AC P01800;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region T957.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RN
RP SEQUENCE.
RX MEDLINE=81216632; PubMed=6787122;
RA Rudikoff S., Potter M.;
RT "Immunoglobulin heavy chains from anti-inulin myeloma proteins:
evidence for a new heavy chain joining segment."
RL J. Immunol. 127:191-194(1981).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; A02073; HVM5AM.
DR HSSP; P01789; 1MCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12691 MW; 7A6D906AAA966E9E CRC64;

Query Match 86.7%; Score 85; DB 1; Length 113;
Best Local Similarity 89.5%; Pred. No. 2.9e-07;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EIRLKSDNYATHYAESVKG 19
Db 50 EIRLKSHNYATHYAESVKG 68

RESULT 9
HV35 MOUSE
ID HV35 MOUSE STANDARD; PRT; 111 AA.
AC P01804;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region HPC76 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=81013937; PubMed=6251474;
RA Bernard O., Gough N.M.;
RT "Nucleotide sequence of immunoglobulin heavy chain joining segments
between translocated VH and mu constant regions genes."
RL Proc. Natl. Acad. Sci. U.S.A. 77:3630-3634(1980).
CC -!- MISCELLANEOUS: THE SEQUENCE OF THE FIRST 197 RESIDUES OF THE C

```

CC REGION WAS ALSO DETERMINED & DIFFERS IN ONLY 3 POSITIONS FROM THE  
 CC CORRESPONDING PORTION OF THE MOUSE MOPC 104E MU CHAIN.

DR PIR; A02074; MHM576.  
 DR HSP; P01789; IMCP.  
 DR InterPro; IPR003006; Ig MHC.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 KW Immunoglobulin V region.  
 FT NON\_TER 111 111  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 12304 MW; 0BDE98EC7348056A CRC64;

Query Match 77.0%; Score 75.5; DB 1; Length 111;  
 Best Local Similarity 89.5%; Pred. No. 1.1e-05;  
 Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 EIRLKSNDYATHYAESVKG 19  
 ||||| ||||| ||||| |||||  
 DB 45 EIRLKS-GYATHYAESVKG 62

RESULT 10  
 HV01\_RAT  
 ID HV01\_RAT STANDARD; PRT; 142 AA.  
 AC P01805;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V region IR2 precursor.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83064537; PubMed=6292865;  
 RA Hellman L., Pettersson U., Engstroem A., Karlsson T., Bennich H.;  
 RT "Structure and evolution of the heavy chain from rat immunoglobulin  
 RT E";  
 RL Nucleic Acids Res. 10:6041-6049(1982).  
 CC -I- MISCELLANEOUS: THE MRNA WAS ISOLATED FROM AN IGE-SECRETING  
 CC IMMUNOCYTOMA THAT ARISES SPONTANEOUSLY IN LOU/C/WSL RATS.

CC PIR; A02075; EVRTR2.  
 DR HSP; P01789; IMCP.  
 DR InterPro; IPR003006; Ig MHC.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 142 IG HEAVY CHAIN V REGION IR2.  
 FT NON\_TER 142 142  
 SQ SEQUENCE 142 AA; 16024 MW; D829B6CFE745DF3B CRC64;

Query Match 58.2%; Score 57; DB 1; Length 142;  
 Best Local Similarity 52.6%; Pred. No. 0.016;  
 Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 EIRLKSNDYATHYAESVKG 19  
 ||||| :||| :|||  
 DB 69 EIRKANNVAVYKSLKG 87

RESULT 11  
 HV21\_MOUSE  
 ID HV21\_MOUSE STANDARD; PRT; 122 AA.  
 AC P01790;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V region MS11.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=81054880; PubMed=6776528;  
 RA Robinson E.A., Appella E.;  
 RA "Complete amino acid sequence of a mouse immunoglobulin alpha chain  
 RT (MOPC 511).";  
 RT Proc. Natl. Acad. Sci. U.S.A. 77:4909-4913(1980).  
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
 CC BINDS PHOSPHORYLCHOLINE.

DR PIR; A02070; AVMS15.  
 DR HSP; P01789; IMCP.  
 DR InterPro; IPR003006; Ig MHC.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 KW Immunoglobulin V region.  
 FT NON\_TER 122 122  
 SQ SEQUENCE 122 AA; 13652 MW; 9F4837731BA50207 CRC64;

Query Match 49.0%; Score 48; DB 1; Length 122;  
 Best Local Similarity 52.9%; Pred. No. 0.43;  
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVKG 19  
 |:::| |:::|  
 DB 52 RNKANDYTTESASVKG 68

RESULT 12  
 HV18\_MOUSE  
 ID HV18\_MOUSE STANDARD; PRT; 123 AA.  
 AC P01787;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V regions TEPC 15/S107/HPCM1/HPCM2/HPCM3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE (TEPC 15).  
 RX MEDLINE=76222762; PubMed=819932;  
 RA Rudikoff S., Potter M.;  
 RA "Size differences among immunoglobulin heavy chains from  
 RT phosphorylcholine-binding proteins";  
 RL Proc. Natl. Acad. Sci. U.S.A. 73:2109-2112(1976).  
 RN [2]  
 RP SEQUENCE FROM N.A. (H107).  
 RX MEDLINE=80199926; PubMed=6769593;  
 RA Early P., Huang H., Davis M., Calame K., Hood L.;  
 RA "An immunoglobulin heavy chain variable region gene is generated from  
 RT three segments of DNA: VH, D and JH.";  
 RL Cell 19:981-992(1980).  
 RN [3]  
 RP SEQUENCE (S107).  
 RX MEDLINE=76110488; PubMed=813561;  
 RA Rudikoff S., Barstad P., Potter M., Hood L.;  
 RA Unpublished results, cited by:  
 RL Hood L., Campbell J.H., Elgin S.C.R.;  
 RL Annu. Rev. Genet. 9:305-353(1975).  
 RN [4]  
 RP SEQUENCE (HPCM1; HPCM2 AND HPCM3).  
 RX MEDLINE=81197602; PubMed=7231520;  
 RA Gearhart P.J., Johnson N.D., Douglas R., Hood L.;  
 RA "IgG antibodies to phosphorylcholine exhibit more diversity than  
 RT their IgM counterparts";  
 RL Nature 291:29-34(1981).  
 CC -I- MISCELLANEOUS: ALL THOSE SEQUENCE APPEARS TO BE IDENTICAL.  
 CC -I- MISCELLANEOUS: THESE CHAINS WERE ISOLATED FROM MYELOMA AND

```

CC HYBRIDOMA PROTEINS THAT BIND PHOSPHORYLCHOLINE.
DR PIR; A02070; AVMS75.
DR HSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Hybridoma.
FT NON_TER 123 123
SQ SEQUENCE 123 AA; 13777 MW; 9D58086DE12F7000 CRC64;

Query Match 49.0%; Score 48; DB 1; Length 123;
Best Local Similarity 52.9%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVKG 19
Db 52 RNKANDYTTTEYSASVKG 68

RESULT 13
HV19 MOUSE
ID HV19 MOUSE STANDARD; PRT; 123 AA.
AC P01788;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region H8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RA Barstad P.;
RL Thesis (1975), California Institute of Technology / Pasadena, U.S.A.
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS PHOSPHORYLCHOLINE.
DR PIR; A02070; AVMS75.
DR HSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT NON_TER 123 123
SQ SEQUENCE 123 AA; 13805 MW; 9D581401912F7000 CRC64;

Query Match 49.0%; Score 48; DB 1; Length 123;
Best Local Similarity 52.9%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVKG 19
Db 52 RNKANDYTTTEYSASVKG 68

RESULT 14
HV22 MOUSE
ID HV22 MOUSE STANDARD; PRT; 123 AA.
AC P01791;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region HPCM6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=81197602; PubMed=7231520;
RA Gearhart P.J.; Johnson N.D.; Douglas R.; Hood L.;
"IGG antibodies to phosphorylcholine exhibit more diversity than
their IGM counterparts.";
RL Nature 291:29-34(1981).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS PHOSPHORYLCHOLINE.
DR PIR; A02070; AVMS75.
DR HSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Hybridoma.
FT NON_TER 123 123
SQ SEQUENCE 123 AA; 13807 MW; A7584FB098B7785D CRC64;

Query Match 49.0%; Score 48; DB 1; Length 123;
Best Local Similarity 52.9%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVKG 19
Db 52 RNKANDYTTTEYSASVKG 68

Search completed: April 22, 2003, 12:51:51
Job time : 10.5 secs

```



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:47:43 ; Search time 38 seconds  
(without alignments)  
103.024 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98

Sequence: 1 EIRKSDNYATHYAEVKG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21.\*

1: sp archaea.\*

2: sp bacteria.\*

3: sp fungi.\*

4: sp human.\*

5: sp invertebrate.\*

6: sp mammal.\*

7: sp mhc.\*

8: sp organelle.\*

9: sp phage.\*

10: sp plant.\*

11: sp rodent.\*

12: sp virus.\*

13: sp vertebrate.\*

14: sp unclassified.\*

15: sp virus.\*

16: sp bacteriap.\*

17: sp archaep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	69.4	64	11 Q61750	Q61750 mus musculus
2	51	52.0	469	11 Q8R3V9	Q8R3V9 mus musculus
3	51	52.0	480	11 Q91XE1	Q91XE1 mus musculus
4	49	50.0	361	2 Q9RBZ2	Q9RBZ2 pseudomonas
5	47	48.0	357	2 Q9RBZ3	Q9RBZ3 pseudomonas
6	46	46.9	2910	10 Q9FND5	Q9FND5 arabidopsis
7	45	45.9	121	4 Q9UL71	Q9UL71 homo sapien
8	45	45.9	422	5 Q8WR61	Q8WR61 lymantria d
9	45	45.9	450	16 Q8UE51	Q8UE51 agrobacteri
10	44	44.9	362	2 Q45962	Q45962 clostridium
11	43	43.9	182	17 Q35037	Q35037 archaeoglob
12	43	43.9	290	17 Q8TZK5	Q8TZK5 pyrococcus
13	43	43.9	310	16 Q99ZF6	Q99ZF6 streptococc
14	43	43.9	474	16 P73837	P73837 synectocyst
15	43	43.9	476	16 Q8YCK5	Q8YCK5 brucella me
16	43	43.9	485	16 Q8YT50	Q8YT50 anabaena sp

17	43	43.9	534	5 Q9XUT8	Q9XUT8 caenorhabdi
18	42.5	43.4	405	2 Q93NC6	Q93NC6 myxococcus
19	42.5	43.4	766	4 P82987	P82987 homo sapien
20	42	42.9	124	12 Q9PZ19	Q9PZ19 xestia c-ni
21	42	42.9	369	16 Q98BT5	Q98BT5 rhizobium l
22	42	42.9	416	5 Q96360	Q96360 hyphantria
23	42	42.9	553	16 Q34751	Q34751 bacillus su
24	42	42.9	679	17 Q59078	Q59078 pyrococcus
25	42	42.9	749	10 Q94F83	Q94F83 oryza sativ
26	42	42.9	1074	5 Q97289	Q97289 plasmodium
27	42	42.9	1261	2 Q69247	Q69247 bacillus li
28	42	42.9	1288	2 Q66071	Q66071 bacillus li
29	42	42.9	1394	5 Q960N9	Q960N9 drosophila
30	42	42.9	1485	5 Q9VJ80	Q9VJ80 drosophila
31	41.5	42.3	751	11 P70535	P70535 rattus norv
32	41	41.8	131	4 Q9UL88	Q9UL88 homo sapien
33	41	41.8	145	5 Q9VA95	Q9VA95 drosophila
34	41	41.8	299	16 Q99UF7	Q99UF7 staphylococ
35	41	41.8	381	11 Q9EQL9	Q9EQL9 rattus norv
36	41	41.8	403	16 Q8YX13	Q8YX13 anabaena sp
37	41	41.8	483	3 Q94206	Q94206 claviceps p
38	41	41.8	532	11 Q92124	Q92124 mus musculu
39	41	41.8	646	16 Q53670	Q53670 mycobacteri
40	41	41.8	656	11 Q920L2	Q920L2 rattus norv
41	41	41.8	661	11 Q921P5	Q921P5 mus musculu
42	41	41.8	663	16 Q9PI96	Q9PI96 campylobact
43	41	41.8	667	17 Q8UOR3	Q8UOR3 pyrococcus
44	41	41.8	724	12 Q9J8C9	Q9J8C9 spodoptera
45	41	41.8	762	16 Q97KV6	Q97KV6 clostridium

## ALIGNMENTS

### RESULT 1

Q61750	Q61750	PRELIMINARY;	PRT;	64 AA.
AC	Q61750;			
DT	01-NOV-1996 (Tremblrel. 01, Created)			
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)			
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)			
DE	JH-Cdelta locus (Fragment).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/C;			
RX	MEDLINE=92017847; PubMed=1922069;			
RA	Owens J.D.Jr. , Finkelman F.D., Mountz J.D., Mushinski J.F.;			
RT	"Nonhomologous recombination at sites within the mouse JH-Cdelta locus			
RT	accompanies Cmu deletion and switch to immunoglobulin D secretion.";			
RL	Mol. Cell. Biol. 11:5660-5670(1991).			
DR	EMBL; M64568; AAA39341.1; --			
DR	HSSP; P01789; IMCP.			
DR	InterPro; IPR003600; IG_like.			
DR	SMART; SM00410; IG_like; 1.			
FT	NON TER			
SQ	SEQUENCE 64 AA; 7594 MW; FE83625079AC2F28 CRC64;			

Query Match 69.4%; Score 68; DB 11; Length 64;  
Best Local Similarity 72.2%; Pred. No. 0.00061;  
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IRLKSDNYATHYAEVKG 19

Db 9 ITVKSNDNYGANYAEVKG 26

### RESULT 2

Q8R3V9 Q8R3V9 PRELIMINARY; PRT; 469 AA.

```

AC Q9R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1; -.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match 52.0%; Score 51; DB 11; Length 469;
Best Local Similarity 55.6%; Pred. No. 4.2;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IRLKSDNYATHYAESVK 19
DB 70 IRLKSDNYATHYAESVK 87

RESULT 3
Q91XE1
ID Q91XE1 PRELIMINARY; PRT; 480 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for IMAGE:4224494) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Tissue=COLON;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010798; AAH10798.1; -.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; ig; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 480 AA; 51936 MW; 20B9234BFF2B41ED CRC64;

Query Match 52.0%; Score 51; DB 11; Length 480;
Best Local Similarity 50.0%; Pred. No. 4.3;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 IRLKSDNYATHYAESVK 19
DB 69 ISNGYATHYFDSMK 84

RESULT 4
Q9RBZ2
ID Q9RBZ2 PRELIMINARY; PRT; 361 AA.
AC Q9RBZ2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE IS970-like transposase.
OS Pseudomonas syringae (pv. glycinea).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=318;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=PG4180;
RX MEDLINE=20015131; PubMed=10545263;
RA Alarcon-Chaidez F.J., Penaloza-Vazquez A., Ullrich M., Bender C.L.;
RT "Characterization of plasmids encoding the phytotoxin coronatine in
  Pseudomonas syringae.";
RL Plasmid 42:210-220(1999).
DR EMBL; AF170066; AAD50977.1; -.
KW Plasmid.
SQ SEQUENCE 361 AA; 41448 MW; 0416AA0203BESA27 CRC64;

Query Match 50.0%; Score 49; DB 2; Length 361;
Best Local Similarity 52.6%; Pred. No. 6.7;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVK 19
DB 247 QLHLVDNYATHKHPKVG 265

RESULT 5
Q9RBZ3
ID Q9RBZ3 PRELIMINARY; PRT; 357 AA.
AC Q9RBZ3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Transposase.
OS Pseudomonas syringae (pv. glycinea).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=318;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20015131; PubMed=10545263;
RA Alarcon-Chaidez F.J., Penaloza-Vazquez A., Ullrich M., Bender C.L.;
RT "Characterization of plasmids encoding the phytotoxin coronatine in
  Pseudomonas syringae.";
RL Plasmid 42:210-220(1999).
DR EMBL; AF169828; AAD50908.1; -.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
KW Plasmid.
SQ SEQUENCE 357 AA; 40339 MW; F21B4028AA5A9BD2 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 357;
Best Local Similarity 58.8%; Pred. No. 14;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 IRLKSDNYATHYAESVK 18
DB 254 IRLKSDNYATHKNDKVK 270

RESULT 6
Q9FND5
ID Q9FND5 PRELIMINARY; PRT; 2910 AA.
AC Q9FND5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similarity to heat shock protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,

```

```

RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones.";
RL DNA Res. 4:291-300(1997).
DR EMBL; AB006702; BAB11602.1; -.
DR EMBL; AB017062; BAB11602.1; JOINED.
SQ SEQUENCE 2910 AA; 325351 MW; A847EC3FE1427DF7 CRC64;

Query Match 46.9%; Score 46; DB 10; Length 2910;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LKSDNYATHYAESVK 18
DB 297 LKEDNHQEEYAESVE 311
|||:|||||:
|||:|||||:

RESULT 7
Q9UL71 PRELIMINARY; PRT; 121 AA.
AC Q9UL71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035043; AAD56279.1; -.
DR HSSP; P01772; 2FBA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF000047; ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 121
FT NON_TER 121
SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFASD50736 CRC64;

Query Match 45.9%; Score 45; DB 4; Length 121;
Best Local Similarity 50.0%; Pred. No. 8.7;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 LKSDNYATHYAESVK 19
DB 51 ISGGGGTYTADSVKG 66
|||:|||||:
|||:|||||:

RESULT 8
Q8WR61 PRELIMINARY; PRT; 422 AA.
AC Q8WR61;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hemolin.
OS Lymantria dispar (Gypsy moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Lymantriidae; Lymantria.
OX NCBI_TaxID=13123;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Denlinger D.L., Lee K.-Y., Horodyski F.M., Valaitis A.P.;
RT "Molecular characterization of the insect immune protein hemolin and
RT the high induction during embryonic diapause in the gypsy moth,
RT Lymantria dispar.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF453868; AAL49765.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF000047; ig; 3.
DR SMART; SM00409; IG; 3.
DR SMART; SM00408; IGC2; 3.
SQ SEQUENCE 422 AA; 47234 MW; 0DC52EC4BF142617 CRC64;

Query Match 45.9%; Score 45; DB 5; Length 422;
Best Local Similarity 53.3%; Pred. No. 37;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 BIRLKSNDYATHYAE 15
DB 42 EVRFKADNYSTALLE 56
|||:|||||:
|||:|||||:

RESULT 9
Q8U651 PRELIMINARY; PRT; 450 AA.
AC Q8U651;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Nitrolicotriacetate monooxygenase, component A.
GN ATU0084 OR AGR_PTI_161.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OG Plasmid pTiC58.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009427; AAL46320.1; -.
DR EMBL; AE007935; AAK91044.1; -.
KW Monooxygenase; Plasmid; Complete proteome.
SQ SEQUENCE 450 AA; 50163 MW; D3ADB8261D68C026 CRC64;

Query Match 45.9%; Score 45; DB 16; Length 450;
Best Local Similarity 58.3%; Pred. No. 40;

```

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 DNYATHYAESVK 18  
 ||| :|||  
 Db 57 DNYGQHFDATVK 68  
 ||| :|||

RESULT 10  
 ID Q45962 PRELIMINARY; PRT; 362 AA.  
 AC Q45962;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE RepB.  
 GN REPB.  
 OS Clostridium butyricum.  
 OG Plasmid pCB101.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1492;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCIB 7423;  
 RA Brehm J.K., Penneck A., Young M., Oultram J.D., Minton N.P.;  
 RT "Physical characterisation of the replication origin of the cryptic  
 RT plasmid pCB101 isolated from Clostridium butyricum NCIB 7423."  
 RL Plasmid 0:0-0(0).  
 DR EMBL; X62684; CAA44562.1; --  
 DR InterPro; IPR000989; Rep.  
 DR Pfam; PF01446; Rep; 1.  
 KW Plasmid.  
 SQ SEQUENCE 362 AA; 43052 MW; FC8A02436E259A49 CRC64;

Query Match 44.9%; Score 44; DB 2; Length 362;  
 Best Local Similarity 53.8%; Pred. No. 45;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ERLKSDNYATHY 13  
 | :|||  
 Db 170 EVTFKNDNYHPHY 182  
 | :|||

RESULT 11  
 ID Q35037 PRELIMINARY; PRT; 182 AA.  
 AC Q35037;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE ISA1083-2, putative transposase.  
 GN AF1352 AND AF2139.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OX NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Klenk H.-P., Clayton R.J., Gwin M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,  
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii H.O., Woese C.R.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon Archaeoglobus fulgidus.";

RL Nature 390:364-370(1997).  
 DR EMBL; AE001010; AAB89893.1; --  
 DR EMBL; AE000956; AAB89112.1; --  
 DR TIGR; AF1352; --  
 DR TIGR; AF2139; --  
 DR InterPro; IPR001584; Rve.  
 DR Pfam; PF00665; Rve; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 182 AA; 20733 MW; 860E5DBDE445CDE7 CRC64;

Query Match 43.9%; Score 43; DB 17; Length 182;  
 Best Local Similarity 58.3%; Pred. No. 30;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 DNYATHYAESVK 18  
 ||| :|||  
 Db 91 DNFKTHAKRVK 102  
 ||| :|||

RESULT 12  
 ID Q87ZK5 PRELIMINARY; PRT; 290 AA.  
 AC Q87ZK5;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical protein PF1986.  
 GN PF1986.  
 OS Pyrococcus furiosus.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=2261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;  
 RT "The complete sequence of the Pyrococcus furiosus genome."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE010292; AAL82110.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 290 AA; 34518 MW; 7697BA58D6794B7E CRC64;

Query Match 43.9%; Score 43; DB 17; Length 290;  
 Best Local Similarity 61.5%; Pred. No. 51;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LKSDNYATHYAES 16  
 ||| :|||  
 Db 34 LKEWYAAHYVDS 46  
 ||| :|||

RESULT 13  
 ID Q99ZF6 PRELIMINARY; PRT; 310 AA.  
 AC Q99ZF6;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative macrolide-efflux protein.  
 GN MREA OR SPY1250.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;  
 RX MEDLINE=21192684; PubMed=11296296;  
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";



Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).  
 ENBL; AEO06564; AAK34105.1; FAD Synth.  
 InterPro; IPRO02606; FAD Synth.  
 Pfam; PFO1687; FAD Synth; 1.  
 ProDom; PD003662; FAD Synth; 1.  
 TIGRFAMs; TIGR00083; FibP; 1.  
 Complete proteome.  
 SEQUENCE 310 AA; 36047 MW; 4A9A7C575CF2B47D C

Query Match 43.9%; Score 43; DB 16; Length 310;  
Best Local Similarity 37.5%; Pred. No. 56;  
Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY	3	RLKSDNYATHYAESVK	18
	:	:	:
D6	103	KVSSDHFITYIKNLK	118

RESULT 14	
P73837	
ID	P73837
AC	PRELIMINARY;
DC	PRT;
DT	474 AA.
DT	01-FEB-1997 (TrEMBLrel. 02, Created)
DT	01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Hypothetical protein slr2004.
DE	SLR2004.
OS	Synechocystis sp. (strain PCC 6803).
OC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX	NCBI_TaxID=1148;
RN	[1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Sugitara M., Sasanoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "sequence analysis of the genome of the unicellular cyanobacterium  
RT *Synechocystis* sp. strain PC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:103-136(1996).  
DR EMBL; D90910; BAA17894.1; -.  
DR InterPro; IPR001258; NHL.  
DR Pfam; PF01436; NHL; 3.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 474 AA; 53482 MW; 52B2F006F34BE7ED CRC64;

Query Match 43.9%; Score 43; DB 16; Length 474;  
Best Local Similarity 42.1%; Pred. NO. 91;  
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 EIRLKSDNYATHYAESVKG 19  
| : : | | | | : |  
Db 87 EVSVKSDAYAEHFLAKPTG 105

[illegible]

RX MEDLINE=20020109; PubMed=11756688;  
 RA Devecchio V.G., Kapatali V., Redkar R.J., Patra G., Mufier C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,  
 RA Jablonowski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,  
 RA Salikov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Haselkorn R., Kyripides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT *Brucella melitensis*,"  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).  
 DR EMBL; AB009688; AAL53765.1; -;  
 DR InterPro; IPR001637; GLnA adenyln.  
 DR InterPro; IPR001691; GLN Synth.  
 DR Pfam; PF00120; gln-synt; 1.  
 DR ProDom; PD001057; GLnA adenyln; 1.  
 KW Ligase; Complete proteome.  
 SQ SEQUENCE 476 AA; 53095 MW; 9DB8DC1C56471313 CRC64;

Query Match 43.9%; Score 43; DB 16; Length 476;  
Best Local Similarity 47.4%; Pred. No. 92;  
Matches 9; Conservative 3; Mismatches 7; Indels

QY 1 EIRKSDNYATHYAESVKG 19  
Db 261 EAALKHDMYATFMAKPIQG 279

Search completed: April 22, 2003, 12:53:52  
Job time : 40 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:45:19 ; Search time 48.1786 Seconds  
(without alignments)  
52.550 Million cell updates/sec

Title: US-09-674-716B-11  
Perfect score: 98  
Sequence: 1 EIRLKSNDYATHYAESVKG 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq 101002:\*

1: /SID22/gcgdata/geneseq/genesecp-emb1/AA1980.DAT:\*

2: /SID22/gcgdata/geneseq/genesecp-emb1/AA1981.DAT:\*

3: /SID22/gcgdata/geneseq/genesecp-emb1/AA1982.DAT:\*

4: /SID22/gcgdata/geneseq/genesecp-emb1/AA1983.DAT:\*

5: /SID22/gcgdata/geneseq/genesecp-emb1/AA1984.DAT:\*

6: /SID22/gcgdata/geneseq/genesecp-emb1/AA1985.DAT:\*

7: /SID22/gcgdata/geneseq/genesecp-emb1/AA1986.DAT:\*

8: /SID22/gcgdata/geneseq/genesecp-emb1/AA1987.DAT:\*

9: /SID22/gcgdata/geneseq/genesecp-emb1/AA1988.DAT:\*

10: /SID22/gcgdata/geneseq/genesecp-emb1/AA1989.DAT:\*

11: /SID22/gcgdata/geneseq/genesecp-emb1/AA1990.DAT:\*

12: /SID22/gcgdata/geneseq/genesecp-emb1/AA1991.DAT:\*

13: /SID22/gcgdata/geneseq/genesecp-emb1/AA1992.DAT:\*

14: /SID22/gcgdata/geneseq/genesecp-emb1/AA1993.DAT:\*

15: /SID22/gcgdata/geneseq/genesecp-emb1/AA1994.DAT:\*

16: /SID22/gcgdata/geneseq/genesecp-emb1/AA1995.DAT:\*

17: /SID22/gcgdata/geneseq/genesecp-emb1/AA1996.DAT:\*

18: /SID22/gcgdata/geneseq/genesecp-emb1/AA1997.DAT:\*

19: /SID22/gcgdata/geneseq/genesecp-emb1/AA1998.DAT:\*

20: /SID22/gcgdata/geneseq/genesecp-emb1/AA1999.DAT:\*

21: /SID22/gcgdata/geneseq/genesecp-emb1/AA2000.DAT:\*

22: /SID22/gcgdata/geneseq/genesecp-emb1/AA2001.DAT:\*

23: /SID22/gcgdata/geneseq/genesecp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	19	21	AA1980.DAT:*
2	98	100.0	137	21	AA1981.DAT:*
3	98	100.0	286	22	AA1982.DAT:*
4	98	100.0	444	21	AA1983.DAT:*
5	95	96.9	100	22	AA1984.DAT:*
6	95	96.9	299	22	AA1985.DAT:*
7	93	94.9	115	14	AA1986.DAT:*
8	93	94.9	116	20	AA1987.DAT:*
9	93	94.9	119	19	AA1988.DAT:*
10	93	94.9	120	18	AA1989.DAT:*

11	93	94.9	122	21	AA1990.DAT:*
12	93	94.9	142	18	AA1991.DAT:*
13	93	94.9	142	20	AA1992.DAT:*
14	93	94.9	143	11	AA1993.DAT:*
15	93	94.9	255	23	AA1994.DAT:*
16	93	94.9	256	23	AA1995.DAT:*
17	93	94.9	503	23	AA1996.DAT:*
18	93	94.9	570	20	AA1997.DAT:*
19	93	94.9	570	20	AA1998.DAT:*
20	91	92.9	151	22	AA1999.DAT:*
21	91	92.9	151	23	AA2000.DAT:*
22	88	89.8	160	16	AA2001.DAT:*
23	88	89.8	554	16	AA2002.DAT:*
24	85	86.7	114	13	AA2003.DAT:*
25	85	86.7	114	13	AA2004.DAT:*
26	85	86.7	114	21	AA2005.DAT:*
27	80.5	82.1	119	18	AA2006.DAT:*
28	80	81.6	17	21	AA2007.DAT:*
29	74	75.5	16	20	AA2008.DAT:*
30	74	75.5	100	22	AA2009.DAT:*
31	74	75.5	119	18	AA2010.DAT:*
32	74	75.5	119	19	AA2011.DAT:*
33	74	75.5	119	22	AA2012.DAT:*
34	74	75.5	119	22	AA2013.DAT:*
35	74	75.5	119	22	AA2014.DAT:*
36	74	75.5	119	22	AA2015.DAT:*
37	74	75.5	119	23	AA2016.DAT:*
38	73	74.5	134	12	AA2017.DAT:*
39	73	74.5	134	12	AA2018.DAT:*
40	73	74.5	139	14	AA2019.DAT:*
41	72	73.5	119	17	AA2020.DAT:*
42	72	73.5	253	16	AA2021.DAT:*
43	71	72.4	119	19	AA2022.DAT:*
44	70	71.4	19	21	AA2023.DAT:*
45	70	71.4	125	21	AA2024.DAT:*

ALIGNMENTS

RESULT 1  
AA1980.DAT:\*

ID AA1980.DAT:\*

AC AA1980.DAT:\*

DT 15-FEB-2000 (first entry)

DE Light chain CDR H2 of mouse anti-CD23 MAb C11.

CD23; FCER1; IGE receptor; monoclonal antibody; C11; mouse; monoclonal antibody; chimeric antibody; humanised antibody; complementarity determining region; CDR; autoimmune disease; inflammation; arthritis; lupus erythematosus; multiple sclerosis; urticaria; nephrotic syndrome; diabetes; uveitis; dermatitis; psoriasis; inflammatory bowel disease; ulcerative colitis; Crohn's disease; Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis; graft-versus-host disease; COPD; bronchitis; diabetes; B-cell malignancy; therapy.

OS Mus musculus.  
XX WO9958679-A1.  
XX 18-NOV-1999.  
XX 07-MAY-1999; 99WO-GB01434.  
XX 09-MAY-1998; 98GB-0009839.  
XX (GLAX ) GLAXO GROUP LTD.  
XX

PI Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 DR WPI; 2000-053101/04.  
 DR N-PSDB; AAZ34743.  
 XX  
 XX Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 XX Claim 1; Page 40; 81pp; English.  
 XX  
 CC This sequence represents complementarity determining region 2  
 CC (CDR H2) of the heavy chain of murine anti-CD23 (FCERII) monoclonal  
 CC antibody C11 (see also AAY32263). The invention provides altered  
 CC antibodies, such as chimeric or humanised antibodies, which comprise  
 CC sufficient of the amino acid sequences of C11 light and heavy chain  
 CC CDRs (see AAY32254-59) to render them capable of binding to the CD23  
 CC type II molecule expressed on haematopoietic cells. The antibodies  
 CC are used to block soluble CD23 formation for treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis,  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX Sequence 19 AA;  
 XX  
 QY 1 EIRLKSNDNYATHYAESVKG 19  
 DB 1 EIRLKSNDNYATHYAESVKG 19  
 XX  
 RESULT 2  
 AAY32260  
 ID AAY32260 standard; Protein; 137 AA.  
 XX  
 AC AAY32260;  
 XX  
 DT 15-FEB-2000 (first entry)  
 XX  
 DE Mouse anti-CD23 MAb C11 heavy chain variable region.  
 XX  
 KW CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse;  
 KW monoclonal antibody; chimeric antibody; humanised antibody;  
 KW complementarity determining region; CDR; autoimmune disease;  
 KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
 KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
 KW urticaria; nephrotic syndrome; glomerulonephritis;  
 KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
 KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
 KW graft-versus-host disease; COPD; bronchitis; diabetes;  
 KW B-cell malignancy; therapy.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 59..63  
 FT /note= "CDR H1"  
 FT Region 78..96  
 FT /note= "CDR H2"  
 FT Region 129..131  
 FT /note= "CDR H3"  
 XX  
 PN W0958679-A1.  
 XX

PD 18-NOV-1999.  
 XX  
 PF 07-MAY-1999; 99WO-GB01434.  
 DR N-PSDB; AAZ34743.  
 XX  
 PR 09-MAY-1998; 98GB-0009839.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 XX WPI; 2000-053101/04.  
 DR N-PSDB; AAZ34743.  
 XX  
 CC Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 XX Claim 8; Fig 1; 81pp; English.  
 XX  
 CC This sequence represents the heavy chain variable region (VH) of  
 CC murine anti-CD23 (FCERII) monoclonal antibody C11. The invention  
 CC provides altered antibodies, such as chimeric or humanised  
 CC antibodies (see AAY32262 and AAY32263), which comprise sufficient of  
 CC the amino acid sequences of the C11 light and heavy chain  
 CC complementarity determining regions (see AAY32254-59) to render them  
 CC capable of binding to the CD23 type II molecule expressed on  
 CC haematopoietic cells. The antibodies are used to block soluble  
 CC CD23 formation in human therapy, for the treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis,  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX Sequence 137 AA;  
 XX  
 QY 1 EIRLKSNDNYATHYAESVKG 19  
 DB 78 EIRLKSNDNYATHYAESVKG 96  
 XX  
 RESULT 3  
 AAB50426  
 ID AAB50426 standard; Protein; 286 AA.  
 XX  
 AC AAB50426;  
 XX  
 DT 13-MAR-2001 (first entry)  
 XX  
 DE Antibody 33F12 catalytic fragment.  
 XX  
 KW Antibody 33F12; ketone compound; antitumour; cytotoxic;  
 KW targeted drug delivery.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200071556-A1.  
 XX  
 PD 30-NOV-2000.  
 XX  
 PF 24-MAY-2000; 2000WO-US14366.  
 XX  
 PR 25-MAY-1999; 99US-0318661.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 XX

Query Match 100.0%; Score 98; DB 21; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-08;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX PI Barbas CF, Shabat D, Rader C, List B, Lerner RA;  
 XX DR WPI; 2001-061339/07.  
 XX DR N-PSDB; AAC90472.  
 XX PT New ketone compounds containing active agents useful as carriers for  
 XX PT e.g. antitumor agents, antibiotics or fluorescent molecules -  
 XX PS Disclosure; Fig 10; 45pp; English.  
 XX CC The present sequence may be used in the activation of new ketone prodrug  
 XX CC compounds containing active agents. The ketone derivatives are useful as  
 XX CC carriers for antitumor agents such as cytotoxic agents, where the  
 XX CC antitumor agent is a microtubule stabilising agent such as paclitaxel,  
 XX CC epothilone or its therapeutically active analogue or an anthracycline  
 XX CC antibiotic such as doxorubicin or its therapeutically active analogue.  
 XX CC The ketone derivatives are useful for targeted drug delivery.  
 XX CC The inactive molecules in the ketone compounds are converted to  
 XX CC active molecules by retro-Michael reaction. The antibody has  
 XX CC bifunctional activity and specifically immunoreacts with cell surface  
 XX CC antigen of a target cell. The active ingredients can be mixed  
 XX CC effectively with excipients as per desired amount along with the  
 XX CC buffering agent to enhance the effectiveness and activity of the  
 XX CC compound.  
 XX SQ Sequence 286 AA;  
 Query Match 100.0%; Score 98; DB 22; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-08;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSDNVATHYAESVKG 19  
 |||||  
 DB 203 EIRLKSDNVATHYAESVKG 221  
 RESULT 4  
 AAY32263  
 ID AAY32263 standard; Protein; 444 AA.  
 XX AC AAY32263;  
 XX DT 15-FEB-2000 (first entry)  
 XX DE Humanised anti-CD23 Mab C11 heavy chain.  
 XX CD23; FCERII; Ige receptor; monoclonal antibody; C11; mouse; human;  
 XX monoclonal antibody; chimeric antibody; humanised antibody;  
 XX complementarity determining region; CDR; autoimmune disease;  
 XX inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
 XX Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
 XX urticaria; nephrotic syndrome; glomerulonephritis;  
 XX inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
 XX Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
 XX graft-versus-host disease; COPD; bronchitis; diabetes;  
 XX B-cell malignancy; therapy.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Region 1..30  
 FT /note= "framework region 1"  
 FT Region 31..35  
 FT /note= "CDR 1"  
 FT Region 36..49  
 FT /note= "framework region 2"  
 FT Region 50..68  
 FT /note= "CDR 2"  
 FT Region 69..100  
 FT /note= "framework region 3"  
 FT Region 101..103

FT Region /note= "CDR 3"  
 FT 104..111  
 FT /note= "framework region 4"  
 FT Region 112..444  
 FT /note= "constant region"  
 XX WO9958679-A1.  
 XX PN 18-NOV-1999.  
 XX PD 07-MAY-1999; 99WO-GB01434.  
 XX PF 09-MAY-1998; 98GB-0009839.  
 XX PR (GLAX ) GLAXO GROUP LTD.  
 XX PA Bonney J, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 XX PI WPI; 2000-053101/04.  
 XX DR N-PSDB; AAZ34748.  
 XX XA Cell receptor specific antibodies useful for treating e.g. arthritis,  
 FT diabetes, multiple sclerosis and psoriasis -  
 XX PS Claim 9; Fig 4; 81pp; English.  
 XX CC This amino acid sequence represents the heavy chain of humanised  
 CC anti-CD23 (FCERII) monoclonal antibody C11, composed of a human  
 CC framework (H5IGKVI) and the heavy chain complementarity  
 CC determining regions (see AAY32257-59) of murine antibody C11. The  
 CC DNA was constructed by splice overlap PCR. The invention provides  
 CC altered antibodies, such as chimeric or humanised antibodies, which  
 CC comprise sufficient of the amino acid sequences of the C11 light  
 CC and heavy chain complementarity determining regions to render them  
 CC capable of binding to the CD23 type II molecule expressed on  
 CC haematopoietic cells. The antibodies are used to block soluble  
 CC CD23 formation in human therapy, for the treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX SQ Sequence 444 AA;  
 Query Match 100.0%; Score 98; DB 21; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSDNVATHYAESVKG 19  
 |||||  
 DB 50 EIRLKSDNVATHYAESVKG 68  
 RESULT 5  
 AAE06973  
 ID AAE06973 standard; Protein; 100 AA.  
 XX AC AAE06973;  
 XX DT 16-OCT-2001 (first entry)  
 XX DE Mouse germline heavy chain variable (VH) region, V(H)22.1.  
 XX XA Mouse; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
 XX neuroprotective; immunosuppressive; human immunodeficiency virus;  
 XX HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 XX inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;

KW multiple sclerosis; atherosclerosis; restenosis; asthma;  
 KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
 KW AIDS; inflammatory glomerulopathy; vascular intervention;  
 KW neointimal hyperplasia; VH; heavy chain variable region.  
 XX  
 OS Mus sp.  
 XX WO200157226-A1.  
 PN 09-AUG-2001.  
 XX 02-FEB-2001; 2001WO-US03537.  
 XX 03-FEB-2000; 2000US-0497625.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
 PI WPI; 2001-488888/53.  
 XX Humanized immunoglobulin for treating a CC-chemokine receptor  
 PT 2-mediated disorder in a patient, comprises a binding specificity for  
 PT CCR2, and a non-human antigen binding region and human immunoglobulin  
 PT -  
 XX Disclosure; Page 152-153; 183pp; English.  
 XX The patent discloses a humanized antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanized antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherosclerosis and atherosclerosis,  
 CC and in the manufacture of a medicament for treating CCR-2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is mouse germline heavy chain variable (VH)  
 CC region, V(H)22.1.  
 XX  
 XX Sequence 100 AA;  
 SQ Query Match 96.9%; Score 95; DB 22; Length 100;  
 Best Local Similarity 94.7%; Pred. No. 3.1e-08;  
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSNDYATHYAESVKG 19  
 :|||||  
 Db 50 QIRLKSNDYATHYAESVKG 68  
 RESULT 6  
 ID AAB50425 standard; Protein; 299 AA.  
 XX AAB50425;  
 AC AAB50425;  
 XX 13-MAR-2001 (first entry)  
 DT Mouse antibody 38C2 catalytic fragment.  
 XX

XX Mouse; antibody 38C2; ketone compound; antitumour; cytotoxic;  
 KW targeted drug delivery.  
 XX Mus sp.  
 PN WO2000071556-A1.  
 XX 30-NOV-2000.  
 XX 24-MAY-2000; 2000WO-US14366.  
 XX 25-MAY-1999; 99US-0318661.  
 XX (SCRI) SCRIPPS RES INST.  
 XX Barbas CF, Shabat D, Rader C, List B, Lerner RA;  
 PI WPI; 2001-061339/07.  
 XX N-PSDB; AAC90471.  
 XX New ketone compounds containing active agents useful as carriers for  
 PT e.g. antitumor agents, antibiotics or fluorescent molecules -  
 XX Disclosure; Fig 9; 45pp; English.  
 XX The present sequence may be used in the activation of new ketone prodrug  
 CC compounds containing active agents. The ketone derivatives are useful as  
 CC carriers for antitumor agents such as cytotoxic agents, where the  
 CC antitumor agent is a microtubule stabilising agent such as paclitaxel,  
 CC epothilone or its therapeutically active analogue or an anthracycline  
 CC antibiotic such as doxorubicin or its therapeutically active analogue.  
 CC The ketone derivatives are useful for targeted drug delivery.  
 CC The inactive molecules in the ketone compounds are converted to  
 CC active molecules by retro-Michael reaction. The antibody has  
 CC bifunctional activity and specifically immunoreacts with cell surface  
 CC antigen of a target cell. The active ingredients can be mixed  
 CC effectively with excipients as per desired amount along with the  
 CC buffering agent to enhance the effectiveness and activity of the  
 CC compound.  
 XX Sequence 299 AA;  
 SQ Query Match 96.9%; Score 95; DB 22; Length 299;  
 Best Local Similarity 94.7%; Pred. No. 1.1e-07;  
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSNDYATHYAESVKG 19  
 :|||||  
 Db 205 EIRLKSNDYATHYAESVKG 223  
 RESULT 7  
 AAR34018  
 ID AAR34018 standard; Protein; 115 AA.  
 XX AAR34018;  
 AC AAR34018;  
 XX 02-AUG-1993 (first entry)  
 DT BW 835 VH.  
 DE Monoclonal antibody; MAb; hybridoma; lung; adenocarcinoma;  
 KW mammary; ovary; prostate; polymorphic epithelial mucin; PEM.  
 XX Synthetic.  
 XX DE4133791-A.  
 XX 15-APR-1993.  
 PD 11-OCT-1991; 91DE-4133791.  
 XX

PR 11-OCT-1991; 91DE-4133791.  
 XX (BEHW ) BEHRINGWERKE AG.  
 XX  
 PI Boslet K, Pfeleiderer P, Seemann G;  
 XX  
 DR WPI; 1993-127068/16.  
 DR N-PSDB; AAQ0046.  
 XX  
 XX New monoclonal antibody BW835 specific for tumour antigens -  
 PT useful for diagnosis and treatment of tumours affecting the  
 PT breasts, ovaries, prostate and lungs  
 XX  
 PS Disclosure; Fig 1a; 24pp; German.  
 XX  
 CC Monoclonal antibody BW 835 is produced by hybridoma cell line BW 835.  
 CC The antibody strongly reacts with lung adenocarcinomas and human  
 CC mammary-, ovary- and prostate carcinomas. It additionally reacts  
 CC with polymorphic epithelial mucin (PEM) but does not react with  
 CC normal human tissue.  
 XX  
 SQ Sequence 115 AA;  
 Query Match 94.9%; Score 93; DB 14; Length 115;  
 Best Local Similarity 94.7%; Pred. No. 7.9e-08;  
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSNDNYATHYAESVKG 19  
 DB 46 EIRLKSNNYATHYAESVKG 64  
 EIRLKSNDNYATHYAESVKG 19  
 EIRLKSNNYATHYAESVKG 64  
 RESULT 8  
 AAY03869  
 ID AAY03869 standard; Protein; 116 AA.  
 AC AAY03869;  
 XX  
 DT 16-JUN-1999 (first entry)  
 XX  
 DE SM3 heavy chain variable region.  
 XX  
 KW SM3 antibody; epitope; mimic; crystal; tumour; MUC1 epitope; allergy;  
 KW immune response; arthritis; multiple sclerosis; asthma; diabetes;  
 KW inflammatory disorder; transplant rejection; graft versus host disease.  
 XX  
 OS Unidentified.  
 XX  
 XX WO9910379-A1.  
 XX  
 XX 04-MAR-1999.  
 XX  
 XX 24-AUG-1998; 98WO-GB02542.  
 XX  
 XX 22-AUG-1997; 97GB-0017946.  
 XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.  
 XX  
 FI Bates PA, Dokurno P, Freemont PS, Snary D, Sternberg MJE;  
 XX  
 DR WPI; 1999-204650/17.  
 DR N-PSDB; AAX31971.  
 XX  
 XX New SM3 antibody crystal structures - used to develop agents for  
 PT treating e.g. tumors, autoimmune disorders, allergies, inflammatory  
 PT disorders or transplant rejection  
 XX  
 XX Disclosure; Page 279-280; 316pp; English.  
 XX  
 CC The invention relates to a process for preparing a crystal using cadmium.  
 CC Structure factors or structural coordinates obtained from the crystal of  
 CC SM3 antibody bound to an epitope can be used to design mimics of the  
 CC antibody or the epitope. The crystals comprise at least an epitope

CC binding fragment of the SM3 antibody bound to a peptide recognised by the  
 CC epitope binding site of SM3. The products and methods can be used to  
 CC develop agents for the detection of tumour cells and for therapy against  
 CC tumours. MUC1 epitope mimics can also be used to prevent or decrease an  
 CC immune response, e.g. in the therapy of diseases caused by autoimmune  
 CC responses (such as arthritis, multiple sclerosis, asthma or diabetes),  
 CC allergies, inflammatory disorders or transplant rejections such as graft  
 CC versus host disease. The present sequence represents the amino acid  
 CC sequence of a heavy chain variable region of SM3 antibody.  
 XX  
 SQ Sequence 116 AA;  
 Query Match 94.9%; Score 93; DB 20; Length 116;  
 Best Local Similarity 94.7%; Pred. No. 7.9e-08;  
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSNDNYATHYAESVKG 19  
 DB 50 EIRLKSNNYATHYAESVKG 68  
 EIRLKSNDNYATHYAESVKG 19  
 EIRLKSNNYATHYAESVKG 68  
 RESULT 9  
 AAW46958  
 ID AAW46958 standard; Protein; 119 AA.  
 AC AAW46958;  
 XX  
 DT 06-JUL-1998 (first entry)  
 XX  
 DE Amino acid sequence of a synthetic branched mucin type glycolipid.  
 XX  
 KW Branched mucin type glycolipid; V region; heavy chain; antibody;  
 KW cancer treatment; diagnosis.  
 XX  
 OS Synthetic.  
 XX  
 XX JP10084963-A.  
 XX  
 PD 07-APR-1998.  
 XX  
 PF 12-SEP-1996; 96JP-0241725.  
 XX  
 PR 12-SEP-1996; 96JP-0241725.  
 XX (TOVJ ) TOSOH CORP.  
 XX  
 DR WPI; 1998-264850/24.  
 DR N-PSDB; AAV22331.  
 XX  
 PT Recognising branched mucin type synthetic glycolipid - using gene  
 PT fragment of an antibody, useful in cancer treatment and diagnosis  
 XX  
 PS Disclosure; Pages 4-5; 6pp; Japanese.  
 XX  
 CC The present sequence represents a branched mucin type synthetic  
 CC glycolipid. A gene fragment encoding the V region of the heavy  
 CC chain of an antibody recognising the present protein is claimed.  
 CC The antibody gene fragment is useful for the development of cancer  
 CC treatments and diagnosing agents.  
 XX  
 SQ Sequence 119 AA;  
 Query Match 94.9%; Score 93; DB 19; Length 119;  
 Best Local Similarity 94.7%; Pred. No. 8.2e-08;  
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSNDNYATHYAESVKG 19  
 DB 50 EIRLKSNNYATHYAESVKG 68  
 EIRLKSNDNYATHYAESVKG 19  
 EIRLKSNNYATHYAESVKG 68  
 RESULT 10  
 AAW01589

KW	specific binding assay; affinity purification; drug targeting; toxin targeting; imaging; genetic; therapeutic.
KX	
OS	Homo sapiens.
XX	
PN	US6054561-A.
XX	
PD	25-APR-2000.
XX	
PF	07-JUN-1995; 95US-0483749.
XX	
XX	21-MAR-1986; 86US-0842476.
PR	08-MAY-1988; 88US-0190778.
PR	08-FEB-1984; 84US-0577976.
PR	11-JAN-1985; 85US-0690750.
PR	11-AUG-1994; 94US-0288981.
XX	
XX	(CHIR ) CHIRON CORP.
XX	
PI	Ring DB;
XX	
DR	WPI; 2000-338508/29.
DR	N-PSDB; AAA38896.
XX	
PT	Monoclonal antibody capable of binding to human breast cancer antigen
PT	useful for affinity purification, drug or toxin targeting, imaging, and
PT	treating cancer -
XX	
XX	Disclosure; Fig 1; 57pp; English.
XX	
PS	The present invention describes a monoclonal antibody (MAB) (I) that
CC	binds to a human breast cancer antigen that is also bound by MAB 454C11
CC	and 520C9 (produced hybridoma ATCC HB8484 and HB8696, respectively). Also
CC	described is a hybridoma that produces (I). (I) is useful in specific
CC	binding assays, affinity purification, drug or toxin targeting, imaging,
CC	and genetic or immunological therapeutics for various cancers. The
CC	present sequence represents a VH domain derived from a 2G3 hybridoma,
CC	which is used in the exemplification of the present invention.
XX	
SQ	Sequence 122 AA;
	Query Match 94.9%; Score 93; DB 21; Length 122;
	Best Local Similarity 94.7%; Pred.No. 8.4e-08;
	Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY	1 EIRLKSNDNYATHYAESVKG 19
	:
Db	50 EIRLKSNNYATHYAESVKG 68
	:
RESULT 12	
AAW06212	
ID	AAW06212 standard; Protein; 142 AA.
AC	
XX	AAW06212;
CC	
DT	12-FEB-1997 (first entry)
XX	
DE	MAB Br-3 heavy chain variable region.
XX	
KW	Chimeric antibody; monoclonal antibody; Br-3; antibody engineering;
KW	tumour; antigen; breast carcinoma; colon carcinoma; lung carcinoma;
KW	ovary carcinoma; cancer; diagnosis; therapy; heavy chain.
XX	
OS	Mus sp.
XX	
XX	US5576184-A.
PN	
PD	19-NOV-1996.
XX	
PF	06-SEP-1988; 88US-0240624.
XX	
XX	06-MAY-1991; 91US-0659401.



```

PR 06-SEP-1988; 88US-0240624.
PR 08-SEP-1988; 88US-0241744.
PR 13-SEP-1988; 88US-0243739.
PR 04-OCT-1988; 88US-0253002.
PR 19-JUN-1989; 89US-0367641.
PR 21-JUL-1989; 89US-0382768.
PR 27-DEC-1994; 94US-0364001.
XX (XOMA ) XOMA CORP.
XX
XX Better MD, Chang CP, Horwitz AH, Lei S, Robinson RR;
XX
XX WPI; 1997-011249/01.
XX N-PSDB; AAT43437.
XX
XX Chimeric mouse-human antibodies - recognise a human tumour antigen,
XX used for the treatment and diagnosis of human cancers
XX
XX Example 3; Fig 15; 102pp; English.
XX
XX The heavy chain variable region (AAW06212) of mouse monoclonal
XX antibody Br-3 is the product of a cDNA clone (AAT43437) isolated
XX from a Br-3 hybridoma cDNA library. MAb Br-3 (IgG1) binds to an
XX antigen that is expressed on the surface of human lung, breast,
XX colon and ovary carcinomas, but not on most normal adult tissues.
XX The heavy chain and light chain variable regions (see also AAW06211)
XX of B38-1 can be linked to human constant regions and expressed in
XX transformed host cells. Novel mouse-human chimeric antibodies (see
XX also AAW06209-10 and AAW06213-18) can be produced that have specificity
XX to human tumour antigens and can be used for the treatment and
XX diagnosis of human cancer.
XX
XX Sequence 142 AA;
SQ
Query Match 94.9%; Score 93; DB 18; Length 142;
Best Local Similarity 94.7%; Pred. No. 1e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRLKSNDNYATHYAESVKG 19
Db 69 EIRLKSNNYATHYAESVKG 87

RESULT 13
AAW85059
ID AAW85059 standard; Protein; 142 AA.
XX
XX AAW85059;
XX
XX 16-APR-1999 (first entry)
XX
XX Mouse Br-3 heavy chain variable region.
XX
XX Heavy chain variable region; murine antibody Br-3; antibody ING-1;
XX chimeric immunoglobulin; human tumour antigen; chimeric antibody;
XX treatment; human cancer.
XX
XX Mus sp.
XX
XX US5843685-A.
XX
XX 01-DEC-1998.
XX
XX 06-JUN-1995; 95US-0466034.
XX
XX 06-SEP-1989; 89WO-US03852.
XX 06-SEP-1988; 88US-0240624.
XX 08-SEP-1988; 88US-0241744.
XX 13-SEP-1988; 88US-0243739.
XX 04-OCT-1988; 88US-0253002.
XX 19-JUN-1989; 89US-0367641.
XX 21-JUL-1989; 89US-0382768.
XX 27-DEC-1994; 94US-0364001.

```

```

PR 06-JUN-1995; 95US-0466034.
XX (XOMA ) XOMA CORP.
XX
XX Better MD, Chang CP, Horwitz AH, Lei S, Robinson RR;
XX
XX WPI; 1999-044574/04.
XX N-PSDB; AAV71155.
XX
XX Chimeric antibody specific for human tumour antigen - useful as
XX immunoassay, imaging or antitumour agent
XX
XX Example 3; Fig 15; 92pp; English.
XX
XX The present sequence represents the heavy chain variable region of
XX murine antibody Br-3. The sequence was used to create chimeric
XX mouse-human immunoglobulins which recognise the human tumour
XX antigen bound by antibody ING-1 (produced by hybridoma cell line
XX ATCC HB 9812). The chimeric antibodies also have an antigen-binding
XX site that competitively inhibits the binding of antibody ING-1, and
XX mediate complement-dependent cytotoxicity of target cells or
XX antibody-dependent cellular cytotoxicity to target cells. The
XX chimeric antibodies can be used for therapeutic purposes in the
XX treatment of human cancer.
XX
XX Sequence 142 AA;
SQ
Query Match 94.9%; Score 93; DB 20; Length 142;
Best Local Similarity 94.7%; Pred. No. 1e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRLKSNDNYATHYAESVKG 19
Db 69 EIRLKSNNYATHYAESVKG 87

RESULT 14
AAW09423
ID AAW09423 standard; Protein; 143 AA.
XX
XX AAW09423;
XX
XX 04-MAR-1993 (first entry)
XX
XX Br-3 Heavy Chain V Region (mouse).
XX
XX Monoclonal antibody; chimera; light; heavy; chain; constant;
XX variable; antigen; diagnosis; cancer; tumour.
XX
XX Mus musculus.
XX
XX WO9002569-A.
XX
XX 22-MAR-1990.
XX
XX 06-SEP-1989; 89WO-US03852.
XX
XX 06-SEP-1988; 88US-0240624.
XX 08-SEP-1988; 88US-0241744.
XX 13-SEP-1988; 88US-0243739.
XX 04-OCT-1988; 88US-0253002.
XX 19-JUN-1989; 89US-0367641.
XX 21-JUL-1989; 89US-0382768.
XX
XX (ITGE-) INT GENETIC ENG INC.
XX
XX Better MD, Horwitz AH, Robinson RR, Lei S, Chang CP;
XX
XX WPI; 1990-115825/15.
XX N-PSDB; AAQ08605.
XX
XX Chimeric mouse-human antibodies - prepd. using genes coding for
XX constant human region murine variable region, esp. to 3 tumour
XX

```

PT antigen  
 PS Claim 13; Page 123 + Fig 15; 173pp; English.  
 CC The sequence is used in the prodn. of a chimeric antibody mol.  
 CC comprising two light chains and two heavy chains, each having a  
 CC constant region (human) and a variable region (murine) having  
 CC specificity to an antigen bound by murine monoclonal antibody  
 CC (Mab) Br-3. The chimeric antibodies can be used for any purpose for  
 CC which the original murine MABs can be used, with the advantage that  
 CC they are more compatible with the human body. They are esp. used for  
 CC the diagnosis and treatment of cancer.  
 XX  
 XX Sequence 143 AA;  
 SQ  
 Query Match 94.9%; Score 93; DB 11; Length 143;  
 Best Local Similarity 94.7%; Pred. No. 1e-07;  
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSNDVYATHYAESVKG 19  
 |||||:|||||||  
 DB 69 EIRLKSNNVYATHYAESVKG 87  
 |||||:|||||||  
 RESULT 15  
 AAU72870  
 ID AAU72870 standard; Protein; 255 AA.  
 XX  
 AC AAU72870;  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX  
 DE P5-23 single chain Fv.  
 XX  
 KW Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;  
 KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;  
 KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;  
 KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;  
 KW sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAPI0;  
 KW helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;  
 KW 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;  
 KW P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;  
 KW P53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200171005-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 26-MAR-2001; 2001WO-EP03414.  
 XX  
 XX 24-MAR-2000; 2000EP-0106467.  
 XX  
 XX (KUFE/) KUFER P.  
 XX  
 XX Kufer P, Riethmuller G, Lutterbuese R, Borschert K, Kischel R;  
 PI Mayer M, Hofmeister R;  
 DR WPI: 2002-055119/07.  
 DR N-PSDB; AAS97144.  
 XX  
 XX Multifunctional polypeptides comprising binding sites that specifically  
 PT recognise extracellular groups of the NKG2D receptor complex and  
 PT domains which function as receptors or ligands, useful for treating  
 PT cancers and infectious diseases -  
 XX  
 XX Example 7; Fig 16; 114pp; English.  
 XX  
 CC The invention relates to a multifunctional polypeptide comprising a  
 CC domain with a binding site that specifically recognises an extracellular  
 CC group of the NKG2D receptor complex and a second domain which functions  
 CC as a receptor or ligand. The polypeptide and its associated

CC polynucleotide are used for the preparation of a pharmaceutical  
 CC composition for the treatment of cancer, infections and/or autoimmune  
 CC conditions. The cancer may be a tumour of the head and neck, stomach,  
 CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,  
 CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,  
 CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.  
 CC The infectious diseases can be caused by viruses, bacteria, fungi,  
 CC protozoa or helminths. The autoimmune diseases include multiple  
 CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior  
 CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent  
 CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and  
 CC autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKG2D  
 CC receptor and the polypeptides of the invention.  
 XX  
 XX Sequence 255 AA;  
 SQ  
 Query Match 94.9%; Score 93; DB 23; Length 255;  
 Best Local Similarity 94.7%; Pred. No. 2e-07;  
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLKSNDVYATHYAESVKG 19  
 |||||:|||||||  
 DB 51 EIRLKSNNVYATHYAESVKG 69  
 |||||:|||||||  
 Search completed: April 22, 2003, 12:51:16  
 Job time : 49.1786 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:53:59 ; Search time 20.0179 Seconds  
(without alignments)  
76.055 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98  
Sequence: 1 EIRLKSNDVATHYAESVKG 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 6: /cgn2\_6/prodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 9: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	285	10	US-09-883-758-4
2	95	96.9	100	10	US-09-840-459-36
3	95	96.9	298	10	US-09-883-758-2
4	91	92.9	151	10	US-09-564-329A-15
5	91	92.9	151	10	US-09-855-153-15
6	91	92.9	151	10	US-09-854-811-15
7	91	92.9	151	10	US-09-934-773-15
8	91	92.9	151	10	US-09-963-620-15
9	74	75.5	119	9	US-10-043-432-5
10	74	75.5	119	10	US-09-756-301A-5
11	74	75.5	119	10	US-09-927-703-5
12	74	75.5	119	10	US-09-766-535A-5
13	74	75.5	119	10	US-09-756-161A-5
14	74	75.5	119	12	US-10-010-229-5
15	74	75.5	119	12	US-10-043-450-5
16	74	75.5	119	12	US-10-044-534-5
17	68	69.4	19	10	US-09-839-666-4
18	68	69.4	27	10	US-09-839-666-21
19	66	67.3	100	10	US-09-840-459-35

20	65	66.3	19	10	US-09-835-087-18	Sequence 18, Appl
21	65	66.3	101	10	US-09-840-459-34	Sequence 34, Appl
22	65	66.3	117	10	US-09-835-087-8	Sequence 8, Appl
23	65	66.3	117	10	US-09-835-087-10	Sequence 10, Appl
24	65	66.3	117	10	US-09-835-087-11	Sequence 11, Appl
25	65	66.3	117	10	US-09-835-087-12	Sequence 12, Appl
26	65	66.3	117	10	US-09-835-087-13	Sequence 13, Appl
27	65	66.3	117	10	US-09-803-739-12	Sequence 12, Appl
28	65	66.3	117	10	US-09-803-739-20	Sequence 20, Appl
29	65	66.3	117	10	US-09-803-739-21	Sequence 21, Appl
30	65	66.3	117	10	US-09-803-739-22	Sequence 22, Appl
31	65	66.3	117	10	US-09-803-739-23	Sequence 23, Appl
32	65	66.3	117	10	US-09-840-459-10	Sequence 10, Appl
33	65	66.3	117	10	US-09-840-459-17	Sequence 17, Appl
34	65	66.3	117	10	US-09-840-459-18	Sequence 18, Appl
35	65	66.3	117	10	US-09-840-459-19	Sequence 19, Appl
36	65	66.3	117	10	US-09-840-459-20	Sequence 20, Appl
37	65	66.3	119	10	US-09-840-459-104	Sequence 104, App
38	65	66.3	148	10	US-09-840-459-100	Sequence 100, App
39	64	65.3	123	10	US-09-893-615-87	Sequence 87, Appl
40	63	64.3	143	10	US-09-881-823-16	Sequence 16, Appl
41	60	61.2	115	10	US-09-840-459-75	Sequence 75, Appl
42	58	59.2	126	10	US-09-840-459-74	Sequence 74, Appl
43	57	58.2	35	9	US-09-956-206A-33	Sequence 33, Appl
44	57	58.2	245	9	US-09-880-748-1979	Sequence 1979, Ap
45	56	57.1	127	10	US-09-840-459-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1  
US-09-883-758-4  
; Sequence 4, Application US/09883758  
; Patent No. US2002058804A1  
; GENERAL INFORMATION:  
; APPLICANT: Barbas III, Carlos F.  
; APPLICANT: Shabat, Doron  
; APPLICANT: Rader, Christoph  
; APPLICANT: List, Benjamin  
; APPLICANT: Lerner, Richard A.  
; TITLE OF INVENTION: PRODRUG ACTIVATION USING CATALYTIC ANTIBODIES  
; FILE REFERENCE: PLF00115  
; CURRENT APPLICATION NUMBER: US/09/883,758  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: US/09/318,661  
; PRIOR FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 285  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Nucleotide  
; OTHER INFORMATION: residue sequence of catalytic fragment  
US-09-883-758-4  
Query Match 100.0%; Score 98; DB 10; Length 285;  
Best Local Similarity 100.0%; Pred. No. 6e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EIRLKSNDVATHYAESVKG 19  
|||  
Db 203 EIRLKSNDVATHYAESVKG 221  
RESULT 2  
US-09-840-459-36  
; Sequence 36, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: LaRosa, Gregory J.



```

; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 09/203,939
; PRIOR FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: 09/251,835
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 09/308,503
; PRIOR FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 151
; TYPE: PRT
; ORGANISM: SCID Mice
US-09-855-153-15

Query Match          92.9%; Score 91; DB 10; Length 151;
Best Local Similarity 89.5%; Pred. No. 3.8e-07;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAEVKG 19
|||||:|||||
DB 69 EIRLSENYATHYAEVKG 87

RESULT 6
US-09-854-811-15
; Sequence 15, Application US/09854811
; Patent No. US20020119157A1
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; APPLICANT: Safran, Douglas C.
; TITLE OF INVENTION: PSCT: PROSTATE STEM CELL ANTIGEN AND USES THEREOF
; FILE REFERENCE: 30435.54US14
; CURRENT APPLICATION NUMBER: US/09/854,811
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 09/564,329
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 09/359,326
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 151
; TYPE: PRT
; ORGANISM: SCID Mice
US-09-854-811-15

Query Match          92.9%; Score 91; DB 10; Length 151;
Best Local Similarity 89.5%; Pred. No. 3.8e-07;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAEVKG 19
|||||:|||||
DB 69 EIRLSENYATHYAEVKG 87

RESULT 7
US-09-934-773-15
; Sequence 15, Application US/09934773
; Patent No. US20020136689A1
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; APPLICANT: Safran, Douglas C.
; TITLE OF INVENTION: PSCT: PROSTATE STEM CELL ANTIGEN AND USES THEREOF
; FILE REFERENCE: 30435.54US14
; CURRENT APPLICATION NUMBER: US/09/934,773
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 09/564,329
; PRIOR FILING DATE: 2000-05-03

```

```

; PRIOR APPLICATION NUMBER: 09/359,326
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 60/113,230
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/120,536
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 60/124,658
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 09/203,939
; PRIOR FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: 09/251,835
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 09/308,503
; PRIOR FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 151
; TYPE: PRT
; ORGANISM: SCID Mice
US-09-934-773-15

Query Match          92.9%; Score 91; DB 10; Length 151;
Best Local Similarity 89.5%; Pred. No. 3.8e-07;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAEVKG 19
|||||:|||||
DB 69 EIRLSENYATHYAEVKG 87

RESULT 8
US-09-963-620-15
; Sequence 15, Application US/09963620
; Patent No. US20020141941A1
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; APPLICANT: Safran, Douglas C.
; TITLE OF INVENTION: PSCT: PROSTATE STEM CELL ANTIGEN AND USES THEREOF
; FILE REFERENCE: 30435.54US14
; CURRENT APPLICATION NUMBER: US/09/963,620
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/564,329
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 09/359,326
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 60/113,230
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/120,536
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 60/124,658
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 09/203,939
; PRIOR FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: 09/251,835
; PRIOR FILING DATE: 1999-02-17

```

;; PRIOR APPLICATION NUMBER: 09/308,503  
;; PRIOR FILING DATE: 1999-05-25  
;; NUMBER OF SEQ ID NOS: 27  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 15  
;; LENGTH: 151  
;; TYPE: PRT  
;; ORGANISM: SCID Mice  
US-09-963-620-15

Query Match 92.9%; Score 91; DB 10; Length 151;  
Best Local Similarity 89.5%; Pred. No. 3.8e-07;  
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19  
Db 69 EIRLSDNYATHYAESVKG 87

RESULT 9  
US-10-043-432-5  
; Sequence 5, Application US/10043432  
; Publication No. US20030054004A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Ghayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; TITLE OF INVENTION: Human Tumor Necrosis Factor  
; FILE REFERENCE: 0975.1005-013  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: 09/927,703  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: U.S. 09/756,398  
; PRIOR FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: U.S. 09/133,119  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: U.S. 08/570,674  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/324,799  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/192,102  
; PRIOR FILING DATE: 1994-10-18  
; PRIOR APPLICATION NUMBER: U.S. 08/192,861  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,093  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/010,406  
; PRIOR FILING DATE: 1993-01-29  
; PRIOR APPLICATION NUMBER: U.S. 07/943,852  
; PRIOR FILING DATE: 1992-09-11  
; PRIOR APPLICATION NUMBER: U.S. 07/853,606  
; PRIOR FILING DATE: 1992-03-18  
; PRIOR APPLICATION NUMBER: U.S. 07/670,827  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Mus Balb/c  
US-10-043-432-5

Query Match 75.5%; Score 74; DB 9; Length 119;  
Best Local Similarity 84.2%; Pred. No. 0.00014;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19  
Db 50 EIRSKINSATHYAESVKG 68

RESULT 10

US-09-756-301A-5  
; Sequence 5, Application US/09756301A  
; Patent No. US20010027249A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Ghayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; TITLE OF INVENTION: Human Tumor Necrosis Factor  
; FILE REFERENCE: 0975.1005-008  
; CURRENT FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: US/09/756,301A  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: U.S. 09/133,119  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/570,674  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/324,799  
; PRIOR FILING DATE: 1994-10-18  
; PRIOR APPLICATION NUMBER: U.S. 08/192,102  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,861  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,093  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/010,406  
; PRIOR FILING DATE: 1993-01-29  
; PRIOR APPLICATION NUMBER: U.S. 08/013,413  
; PRIOR FILING DATE: 1993-02-02  
; PRIOR APPLICATION NUMBER: U.S. 07/943,852  
; PRIOR FILING DATE: 1992-09-11  
; PRIOR APPLICATION NUMBER: U.S. 07/853,606  
; PRIOR FILING DATE: 1992-03-18  
; PRIOR APPLICATION NUMBER: U.S. 07/670,827  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Mus Balb/c  
US-09-756-301A-5

Query Match 75.5%; Score 74; DB 10; Length 119;  
Best Local Similarity 84.2%; Pred. No. 0.00014;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EIRLSDNYATHYAESVKG 19  
Db 50 EIRSKINSATHYAESVKG 68

RESULT 11

US-09-927-703-5  
; Sequence 5, Application US/09927703  
; Patent No. US20020022720A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Ghayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; TITLE OF INVENTION: Human Tumor Necrosis Factor



RESULT 15  
US-10-043-450-5  
; Sequence 5, Application US/10043450  
; Patent No. US20020141996A1  
; GENERAL INFORMATION:  
; APPLICANT: Le, Junming  
; APPLICANT: Vilcek, Jan  
; APPLICANT: Daddona, Peter  
; APPLICANT: Ghayeb, John  
; APPLICANT: Knight, David M.  
; APPLICANT: Siegel, Scott  
; TITLE OF INVENTION: Anti-TNF Antibodies and Peptides of  
; TITLE OF INVENTION: Human Tumor Necrosis Factor  
; FILE REFERENCE: 0975.1005-013  
; CURRENT APPLICATION NUMBER: US/10/043,450  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: 09/927,703  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: U.S. 09/756,398  
; PRIOR FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: U.S. 09/133,119  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: U.S. 08/570,674  
; PRIOR FILING DATE: 1995-12-11  
; PRIOR APPLICATION NUMBER: U.S. 08/324,799  
; PRIOR FILING DATE: 1994-10-18  
; PRIOR APPLICATION NUMBER: U.S. 08/192,102  
; PRIOR FILING DATE: 1994-02-04  
; PRIOR APPLICATION NUMBER: U.S. 08/192,861  
; PRIOR FILING DATE: 1994-02-04



GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:07:14 ; Search time 15 Seconds  
(without alignments)  
102.543 Million cell updates/sec

Title: US-09-674-716B-3  
Perfect score: 81  
Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues 2770  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	33.3	15	2 C44101	calmodulin, vasoac
2	27	33.3	16	2 S09732	photosystem I prot
3	26	32.1	12	2 S09082	proteasome chain 1
4	25	30.9	13	2 S09733	photosystem I prot
5	22	27.2	10	2 B24736	inhibin beta-B cha
6	22	27.2	13	2 S54344	glyceraldehyde-3-p
7	22	27.2	14	2 A28018	very late antigen-
8	22	27.2	15	2 S21240	alpha-glucosidase
9	22	27.2	15	2 S21202	glucan 1,4-alpha-g
10	22	27.2	16	2 A44101	calmodulin, vasoac
11	21	25.9	11	2 PH0914	T-cell receptor be
12	21	25.9	14	2 PC2373	probable IMP dehyd
13	21	25.9	14	2 PH1311	Ig heavy chain DJ
14	21	25.9	15	2 A47146	topoisomerase I -
15	21	25.9	15	2 A36279	chemoattractant pr
16	21	25.9	15	2 B26501	lipoprotein lipase
17	20	24.7	10	2 PT0243	Ig heavy chain CRD
18	20	24.7	13	2 JQ2309	hypothetical 1.6K
19	20	24.7	13	2 JQ2319	hypothetical 1.6K
20	20	24.7	15	2 I46512	troponin - rabbit
21	20	24.7	15	2 JF0101	fibrinogen alpha c
22	20	24.7	15	2 PH1314	Ig heavy chain DJ
23	20	24.7	15	2 PH1616	Ig H chain V-D-J r
24	20	24.7	15	2 S71396	dihydropyrimidine
25	20	24.7	16	2 C39509	mannose-specific 1
26	19.5	24.1	16	2 E28027	protein p8 - curle
27	19	23.5	10	2 A60476	S-layer protein -
28	19	23.5	11	2 A29806	acidic proline-ric
29	19	23.5	12	2 S68402	NAD(+)-glycohydrol

30	19	23.5	12	2 PH0930	T-cell receptor be
31	19	23.5	13	2 G83988	hypothetical prote
32	19	23.5	16	2 G45681	orf 61.1 - phage T
33	18	22.2	8	2 A28719	thymic humoral fac
34	18	22.2	9	2 B20569	serum amyloid P-co
35	18	22.2	10	2 S10926	inhibin beta-A cha
36	18	22.2	10	2 G58501	48K bile/gallblad
37	18	22.2	11	2 B29806	acidic proline-ric
38	18	22.2	11	2 PD0441	translation elonga
39	18	22.2	13	2 B26093	microbial collagen
40	18	22.2	13	2 S47390	T-cell antigen rec
41	18	22.2	14	1 NTKN1M	alpha-conotoxin MI
42	18	22.2	14	2 E33098	214K exoantigen (v
43	18	22.2	15	2 S32677	nitrogenase cofact
44	18	22.2	15	2 PA0058	protein QF200022 -
45	18	22.2	15	2 G35141	T-cell receptor de

ALIGNMENTS

RESULT 1

C44101  
calmodulin, vasoactive intestinal peptide-binding protein, VIP binding protein, p18 - gu  
C;Species: Cavia porcellus (guinea pig)  
C;Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997  
C;Accession: C44101  
R;Stallwood, D.; Brugger, C.H.; Baggenstoss, B.A.; Stemmer, P.M.; Shiraga, H.; Landers,  
J. Biol. Chem. 267, 19617-19621, 1992  
A;Title: Identity of a membrane-bound vasoactive intestinal peptide-binding protein with  
A;Reference number: A44101; MUID:92406918; PMID:1527080  
A;Accession: C44101  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-15 <STA>  
A;Experimental source: lung membranes  
A;Note: sequence extracted from NCBI backbone (NCBIP:114109)  
C;Keywords: intestine

Query Match 33.3%; Score 27; DB 2; Length 15;  
Best Local Similarity 50.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 KDGKTYLN 16  
||| |:  
Db 4 KDGNYIS 11

RESULT 2

S09732  
photosystem I protein psaj - spinach chloroplast (fragment)  
C;Species: chloroplast Spinacia oleracea (spinach)  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 16-Feb-1997  
C;Accession: S09732  
R;Ikeuchi, M.; Hirano, A.; Hiyama, T.; Inoue, Y.  
FEBS Lett. 263, 274-278, 1990  
A;Title: Polypeptide composition of higher plant photosystem I complex. Identification c  
A;Reference number: S09730; MUID:92042987; PMID:2185953  
A;Accession: S09732  
A;Molecule type: protein  
A;Residues: 1-16 <IKE>  
C;Genetics:  
A;Gene: psaj  
A;Genome: chloroplast  
C;Keywords: chloroplast; photosynthesis; photosystem I; transmembrane protein

Query Match 33.3%; Score 27; DB 2; Length 16;  
Best Local Similarity 62.5%; Pred. No. 4.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 KDGKTYLN 16  
||| |:  
Db 2 RDKFYLS 9

## RESULT 3

S09082  
proteasome chain 1 - rat (fragment)  
N;Alternate names: multicatalytic proteinase chain 1  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 21-Aug-1998  
C;Accession: S09082  
R;Lilley, K.S.; Davison, M.D.; Rivett, A.J.

FESB Lett. 262, 327-329, 1990  
A;Title: N-terminal sequence similarities between components of the multicatalytic protease  
A;Reference number: S09082; MUID:90242957; PMID:2335214

A;Accession: S09082

A;Molecule type: protein

A;Residues: 1-12 <LIL>

C;Superfamily: multicatalytic endopeptidase, complex chain C9

Query Match 32.1%; Score 26; DB 2; Length 12;

Best Local Similarity 66.7%; Pred. No. 4.4e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLYKDG 11

Db :|||

6 VVYKDG 11

## RESULT 4

S09733  
photosystem I protein psaJ - garden pea chloroplast (fragment)  
C;Species: chloroplast Pisum sativum (garden pea)  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 16-Feb-1997  
C;Accession: S09733  
R;Kneuch, M.; Hirano, A.; Hiyama, T.; Inoue, Y.

FESB Lett. 263, 274-278, 1990

A;Title: Polypeptide composition of higher plant photosystem I complex. Identification of

A;Reference number: S09730; MUID:90242987; PMID:2185953

A;Accession: S09733

A;Molecule type: protein

A;Residues: 1-13 <LKE>

C;Genetics:

A;Gene: psaJ

A;Genome: chloroplast

C;Keywords: chloroplast; photosynthesis; photosystem I; transmembrane protein

Query Match 30.9%; Score 25; DB 2; Length 13;

Best Local Similarity 71.4%; Pred. No. 7.1e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 9 KDKKTYL 15

Db :|||

2 RDLKTYL 8

## RESULT 5

B24736  
inhibin beta-B chain - pig (fragment)  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 28-Sep-1987 #sequence\_revision 28-Sep-1987 #text\_change 30-Sep-1993  
C;Accession: B24736  
R;King, N.; Ying, S.Y.; Ueno, N.; Esch, F.; Denoroy, L.; Guillemin, R.  
Proc. Natl. Acad. Sci. U.S.A. 82, 7217-7221, 1985

A;Reference number: A24736; MUID:86042637; PMID:3864157

A;Accession: B24736

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <LIN>

Query Match 27.2%; Score 22; DB 2; Length 10;

Best Local Similarity 55.6%; Pred. No. 1.7e+03;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 LYDKGTYL 15

Db :|||

2 LEXDGTNL 10

## RESULT 6

S54344  
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - bovine (fragment)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 27-Oct-1995 #sequence\_revision 30-Jan-1998 #text\_change 03-Jun-2002  
C;Accession: S54344

R;Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.

Biochem. J. 306, 551-555, 1995

A;Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia

A;Reference number: S54343; MUID:95194333; PMID:7887910

A;Accession: S54344

A;Molecule type: protein

A;Residues: 1-13 <OKA>

C;Keywords: oxidoreductase

Query Match 27.2%; Score 22; DB 2; Length 13;

Best Local Similarity 50.0%; Pred. No. 2.2e+03;

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 7 LYXDGK 12

Db :|||

7 LWRDGR 12

## RESULT 7

A28018  
very late antigen-1 alpha chain - human (fragment)  
N;Alternate names: VLA-1 alpha chain  
C;Species: Homo sapiens (man)  
C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 12-May-1994  
C;Accession: A28018  
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.  
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987

A;Title: The very late antigen family of heterodimers is part of a superfamily of molecu

A;Reference number: A94151; MUID:87204112; PMID:3033641

A;Accession: A28018

A;Molecule type: protein

A;Residues: 1-14 <TAK>

C;Keywords: duplication; heterodimer; membrane protein

Query Match 27.2%; Score 22; DB 2; Length 14;

Best Local Similarity 57.1%; Pred. No. 2.4e+03;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 9 KDKKTYL 15

Db :|||

6 KDSMTFL 12

## RESULT 8

S21240  
alpha-glucosidase (EC 3.2.1.20) I - Bacillus "thermoamyloliquefaciens" (fragment)  
C;Species: Bacillus "thermoamyloliquefaciens"  
C;Date: 22-Nov-1993 #sequence\_revision 01-Nov-1996 #text\_change 21-Aug-1998  
C;Accession: S21240  
R;Suzuki, Y.; Yonezawa, K.; Hattori, M.; Takii, Y.  
Eur. J. Biochem. 205, 249-256, 1992

A;Title: Assignment of Bacillus thermoamyloliquefaciens KP1071 alpha-glucosidase I to an

ence and in structural parameters calculated from the amino acid composition.

A;Reference number: S21202; MUID:92209510; PMID:1555585

A;Accession: S21240

A;Molecule type: protein

A;Residues: 1-15 <SUZ>

A;Experimental source: strain KP1071

C;Superfamily: alpha-glucosidase; alpha-amylase core homology

C;Keywords: glycosidase; hydrolase

Query Match 27.2%; Score 22; DB 2; Length 15;

Best Local Similarity 36.4%; Pred. No. 2.6e+03;

Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KSLLYKDGKTY 14  
| : | : |  
Db 2 KKAHWKEGVVY 12

RESULT 9  
S21202  
Glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Bacillus stearothermophilus (fragment)  
N:Alternate names: exo-alpha-1-4-glucosidase I  
C:Species: Bacillus stearothermophilus  
C:Date: 22-Nov-1993 #sequence\_revision 01-Nov-1996 #text\_change 21-Aug-1998  
C:Accession: S21202  
R:Suzuki, Y.; Yonezawa, K.; Hattori, M.; Takii, Y.  
A:Title: Assignment of Bacillus thermoamylioliquefaciens KPI071 alpha-glucosidase I to an  
nase and in structural parameters calculated from the amino acid composition.  
A:Reference number: S21202; MUID:92209510; PMID:1555585  
A:Accession: S21202  
A:Molecule type: protein  
A:Residues: 1-15 <SUZ>  
A:Experimental source: ATCC 12016  
C:Superfamily: alpha-glucosidase; alpha-amylase core homology  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 27.2%; Score 22; DB 2; Length 15;  
Best Local Similarity 36.4%; Pred. No. 2.6e+03;  
Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KSLLYKDGKTY 14  
| : | : |  
Db 2 KKTWKEGVAY 12

RESULT 10  
A44101  
calmodulin, vasoactive intestinal peptide-binding protein, VIP binding protein, p18 - gu  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 23-Sep-1999  
C:Accession: A44101  
R:Stallwood, D.; Brugger, C.H.; Baggenstoss, B.A.; Stemmer, P.M.; Shiraga, H.; Landers,  
J. Biol. Chem. 267, 19617-19621, 1992  
A:Title: Identity of a membrane-bound vasoactive intestinal peptide-binding protein with  
A:Reference number: A44101; MUID:92406918; PMID:1527080  
A:Accession: A44101  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-16 <STA>  
A:Experimental source: lung membranes  
A:Note: sequence extracted from NCBI backbone (NCBI:114120)  
C:Superfamily: calmodulin; calmodulin repeat homology  
C:Keywords: EF hand; intestine

Query Match 27.2%; Score 22; DB 2; Length 16;  
Best Local Similarity 71.4%; Pred. No. 2.7e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SLLYKDG 11  
| : | : |  
Db 4 SLFDKDG 10

RESULT 11  
PH0914  
T-cell receptor beta chain V-D-J region (isolate 7) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C:Accession: PH0914  
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenberg, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy  
A:Reference number: PH0914; MUID:92078857; PMID:1836012

Accession: PH0914  
Molecule type: mRNA  
Residues: 1-11 <GOL>  
Experimental source: myelin basic protein-immunized lymph node  
Keywords: T-cell receptor

Query Match 25.9%; Score 21; DB 2; Length 11;  
Best Local Similarity 62.5%; Pred. No. 2.8e+03;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSKSLLYK 9  
| : | : |  
Db 3 SSDSLSYE 10

RESULT 12  
PC2373  
probable IMP dehydrogenase (EC 1.1.1.205) [similarity] - Bacillus cereus (strain ts-4)  
C:Species: Bacillus cereus  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 02-Sep-2000  
C:Accession: PC2373  
R:Matsuno, K.; Miyamoto, T.; Yamaguchi, K.; Saved, M.A.; Kajiwara, T.; Hatano, S.  
Biosci. Biotechnol. Biochem. 59, 231-235, 1995  
A:Title: Identification of DNA-binding proteins changed after induction of sporulation i  
A:Reference number: PC2369; MUID:95218265; PMID:7766022  
A:Accession: PC2373  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14 <MAS>  
C:Superfamily: IMP dehydrogenase; CBS homology; IMP dehydrogenase amino-terminal homology  
C:Keywords: GMP biosynthesis; NAD; oxidoreductase; purine nucleotide biosynthesis

Query Match 25.9%; Score 21; DB 2; Length 14;  
Best Local Similarity 36.4%; Pred. No. 3.5e+03;  
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 KSLLYKDGKTY 14  
| : | : |  
Db 3 ESKFVKEGLTF 13

RESULT 13  
PH1311  
Ig heavy chain DJ region (clone C68-101) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: PH1311  
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.  
J. Exp. Med. 176, 1577-1581, 1992  
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph  
A:Reference number: PH1302; MUID:93094761; PMID:1460419  
A:Accession: PH1311  
A:Molecule type: DNA  
A:Residues: 1-14 <WAS>  
C:Keywords: heterotetramer; immunoglobulin

Query Match 25.9%; Score 21; DB 2; Length 14;  
Best Local Similarity 50.0%; Pred. No. 3.5e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 LLYKDGKT 13  
| : | : |  
Db 2 ILYHPGTT 9

RESULT 14  
A47146  
topoisomerase I - vaccinia virus (fragment)  
C:Species: vaccinia virus  
C:Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 31-Oct-1997  
C:Accession: A47146  
R:Klempner, N.; Traktman, P.  
J. Biol. Chem. 269, 15887-15899, 1994

A:Title: Biochemical analysis of mutant alleles of the vaccinia virus topoisomerase I ca  
A:Reference number: A47146; MUID:93340198; PMID:8393454

A:Accession: A47146

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-15 <KLE>

A:Cross-references: GB:L13447

C:Superfamily: vaccinia virus DNA topoisomerase

Query Match 25.9%; Score 21; DB 2; Length 15;  
Best Local Similarity 66.7%; Pred. No. 3.7e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 9 KDGKTY 14

|||

Db 5 KDLRTY 10

#### RESULT 15

A36279

chemoattractant protein - earthworm (Lumbricus terrestris) (fragment)

C:Species: Lumbricus terrestris (common earthworm)

C:Date: 18-Jan-1991 #sequence\_revision 18-Jan-1991 #text\_change 30-Sep-1993

C:Accession: A36279

R:Jiang, X.C.; Inouchi, J.; Wang, D.; Halpern, M.

J. Biol. Chem. 265, 8736-8744, 1990

A:Title: Purification and characterization of a chemoattractant from electric shock-indu  
snakes.

A:Reference number: A36279; MUID:90256600; PMID:2160465

A:Accession: A36279

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-15 <JIA>

Query Match 25.9%; Score 21; DB 2; Length 15;  
Best Local Similarity 45.5%; Pred. No. 3.7e+03;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 SLLYKDGKTYL 15

|||

Db 1 ALVCPFGFTYL 11

Search completed: April 22, 2003, 13:13:24

Job time : 16 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:55:49 ; Search time 24 Seconds  
(without alignments)  
27.651 Million cell updates/sec

Title: US-09-674-716B-3  
Perfect score: 81  
Sequence: 1 RSSKSLLYKDKTYLNL 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues 822  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwisProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	25	30.9	13	PSAJ_PEA	P17229 pisum sativ
2	22	27.2	15	MALT_BACTQ	P80072 bacillus th
3	21	25.9	15	MM01_RAT	P81563 rattus norv
4	20	24.7	14	MY14_BISFO	P46979 eisenia foe
5	20	24.7	15	FIBA_NAPL	P12801 anas platyr
6	19	23.5	10	SLAP_BACTG	P43325 bacillus th
7	18	22.2	9	SAMP_MUSCA	P19095 mustelus ca
8	18	22.2	10	PPCK_FASHE	P80525 fasciola he
9	18	22.2	10	SPI_HALRO	Q10997 halocynthia
10	18	22.2	14	CXA1_CONMA	P01521 conus magus
11	18	22.2	14	EFTU_CANFA	P54835 canis famil
12	18	22.2	14	MCRZ_METTM	P58816 methanobact
13	18	22.2	15	NIS1_ANAVA	Q44507 anabaena va
14	17	21.0	10	GONI_PETMA	P04378 petromyzon
15	17	21.0	10	SP34_DICMU	P81545 dictyosteli
16	17	21.0	10	TEMK_RANTE	P56923 rana tempor
17	17	21.0	10	UXA6_CHLTR	P38007 chlamydia t
18	17	21.0	11	CSU5_BACSU	P81095 bacillus su
19	17	21.0	11	TKN2_UFERU	P08616 uperoeteia r
20	17	21.0	12	V14K_WSSV	P82006 white spot
21	17	21.0	13	LIGA_TRAVE	P20011 trametes ve
22	17	21.0	14	MY14_PHEVI	P46980 pheretima v
23	17	21.0	14	NEJ2_FASHE	P80526 fasciola he
24	17	21.0	14	UN07_CLOPA	P81352 clostridium
25	17	21.0	15	ATP2_PINPS	P81663 pinus pinas
26	17	21.0	15	CIQA_RAT	P31720 rattus norv
27	17	21.0	15	PC20_BRANA	P81096 brassica na
28	17	21.0	15	UC14_MAIZE	P80620 zea mays (m
29	17	21.0	16	MLB_SQUAC	P01207 squalus aca
30	16	19.8	10	COXO_RAT	P80432 rattus norv
31	16	19.8	10	TKNE_RANCA	P22689 rana catesb
32	16	19.8	10	TP1S_NICPL	P19118 nicotiana p
33	16	19.8	11	MORN_HUMAN	P01163 homo sapien

34	16	19.8	14	1	CXA1_CONCN	P56973 conus conso
35	16	19.8	14	1	RS19_PPWBP	Q52093 pigeon pea
36	16	19.8	15	1	MILT_ONCKE	P81037 oncorhynch
37	15	18.5	9	1	OXYA_SCYCA	P42996 scyllorhinu
38	15	18.5	9	1	OXYA_SQUAC	P42999 squalus aca
39	15	18.5	10	1	SYK_CAMUP	Q46464 campylobact
40	15	18.5	13	1	LMT4_LOCM1	P41490 locusta mig
41	15	18.5	13	1	MPI_MICOC	P81532 microplitis
42	15	18.5	15	1	CXA1_CONGE	P01519 conus geogr
43	15	18.5	15	1	FGF1_CANFA	P18651 canis famil
44	15	18.5	15	1	PGKH_PHYPA	P80659 physcomitre
45	15	18.5	15	1	UC19_MAIZE	P80625 zea mays (m

ALIGNMENTS

RESULT 1  
PSAJ\_PEA  
ID PSAJ\_PEA STANDARD; PRT; 13 AA.  
AC P17229;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Photosystem I reaction center subunit IX (PSI-J) (Fragment).  
GN PSAJ.  
OS Pisum sativum (Garden pea).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=90242987; PubMed=2185953;  
RA Ikeuchi M., Hirano A., Hiyama T., Inoue Y.;  
RT "Polypeptide composition of higher plant photosystem I complex.  
RT Identification of psal, psaj and psak gene products.";  
RL FEBS Lett. 263:274-278(1990).  
CC -1- FUNCTION: May help in the organization of the psae and psaf  
CC subunits.  
CC -1- SIMILARITY: BELONGS TO THE PSAJ FAMILY.  
CC PIR; S09733; S09733.  
DR Chloroplast; Photosystem I; Photosynthesis; Transmembrane.  
FT TRANSMEM 7 >13 POTENTIAL.  
FT NON\_TER 13  
SQ SEQUENCE 13 AA; 1516 MW; 9E2E45D11FDE3B41 CRC64;  
Query Match 30.9%; Score 25; DB 1; Length 13;  
Best Local Similarity 71.4%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 9 KDKTYL 15  
DB 2 RDLKTYL 8  
RESULT 2  
MALT\_BACTQ  
ID MALT\_BACTQ STANDARD; PRT; 15 AA.  
AC P80072;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-DEC-1992 (Rel. 24, Last annotation update)  
DE Maltase (EC 3.2.1.20) (Alpha-glucosidase I) (Fragment).  
OS Bacillus thermoamyloliquefaciens.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1425;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=KPI071 / FERM P8477;  
RX MEDLINE=92209510; PubMed=1555585;  
RA Suzuki Y., Yonezawa K., Hattori M., Takii Y.;

```

RT "Assignment of Bacillus thermoamylioliquefaciens KP1071
RT alpha-glucosidase I to an exo-alpha-1,4-glucosidase, and its striking
RT similarity to bacillary oligo-1,6-glucosidases in N-terminal sequence
RT and in structural parameters calculated from the amino acid
RT composition."
RL Eur. J. Biochem. 205:249-256(1992).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
CC linked D-glucose residues with release of D-glucose.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR PIR; S21240; S21240.
KW Hydrolase; Glycosidase.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1929 MW; 62B4CE501F2D3042 CRC64;

Query Match 27.2%; Score 22; DB 1; Length 15;
Best Local Similarity 36.4%; Pred. No. 6.6e+02;
Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KSLYKDGKTY 14
DB 2 KKAWKEGVVY 12

RESULT 3
MM01_RAT STANDARD; PRT; 15 AA.
AC P81563;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interstitial collagenase (EC 3.4.24.7) (Matrix metalloproteinase-1)
DE (MMP-1) (Fibroblast collagenase) (Myocardial collagenase) (fragment).
GN MMP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX MEDLINE=96201136; PubMed=8605638;
RT Tyagi S.C.; Cleutjens J.P.M.;
RT "Myocardial collagenase: purification and structural
RT characterization."
RL Can. J. Cardiol. 12:165-171(1996).
CC -1- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN
CC THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X. MAY
CC PLAY A ROLE IN THE DETERIORATION OF THE HEART WALL EXTRACELLULAR
CC MATRIX PROTEINS DURING THE ONSET OF DIALATED CARDIOMYOPATHY.
CC -1- CATALYTIC ACTIVITY: Cleaves preferentially one bond in native
CC collagen. Cleavage of the triple helix of collagen at about three-
CC quarters of the length of the molecule from the N-terminus, at
CC 775-Gly-Ile-776 in the alpha-1(I) chain. Cleaves synthetic
CC substrates and alpha-macroglobulins at bonds where P1' is a
CC hydrophobic residue.
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -1- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE
CC ACTIVATION PEPTIDE.
CC -1- PTM: THE N-TERMINAL IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Collagen degradation;
KW Extracellular matrix.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1787 MW; 15A57D24C0F6FD80 CRC64;

Query Match 25.9%; Score 21; DB 1; Length 15;
Best Local Similarity 40.0%; Pred. No. 9.8e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RSSKSLVKD 10
:|:|:|

```

Db 4 KSEKNADFKD 13

## RESULT 4

```

MY14_EISFO STANDARD; PRT; 14 AA.
AC P46979;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Myoactive tetradecapeptide (ETP).
OS Eisenia foetida (Common brandling worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
OX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Gut;
RX MEDLINE=96087879; PubMed=8532604;
RA Ukena K.; Umi T.; Matsushima O.; Ikeda T.; Fujita T.; Minakata H.;
RA Nomoto K.;
RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia
RT foetida."
RL Peptides 16:995-999(1995).
CC -1- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT
CC MUSCLES.
CC -1- SIMILARITY: TO INSECTS ALLATOTROPIN.
CC KW Neuropeptide; Amidation.
FT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1478 MW; CC9ABEF941CD91AD CRC64;

Query Match 24.7%; Score 20; DB 1; Length 14;
Best Local Similarity 75.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 YKDG 11
:|:|:|
DB 2 FKDG 5

```

## RESULT 5

```

FIBA_ANAPL STANDARD; PRT; 15 AA.
AC P12601;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE.
RX MEDLINE=85168193; PubMed=3983613;
RA Min Y.; Ping Z.; Yaoshi Z.;
RT "Purification and primary structures of duck fibrinopeptides A and
RT B."
RL Sci. Sin., B, Chem. Biol. Agric. Med. Earth Sci. 28:31-35(1985).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR PIR; JP0101; JP0101.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 15
FT MOD_RES 1 1
FIBRINOPEPTIDE A.
PYRROLIDONE CARBOXYLIC ACID.

```

FT NON TER 15 15  
SQ SEQUENCE 15 AA; 1580 MW; D78A51FF88B40373 CRC64;

Query Match 24.7%; Score 20; DB 1; Length 15;  
Best Local Similarity 36.4%; Pred. No. 1.5e+03;  
Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RSKSLLYKDG 11  
: |||:  
Db 1 QDGKSSFOKEG 11

RESULT 6  
SLAP\_BACTG  
ID SLAP\_BACTG STANDARD; PRT; 10 AA.  
AC P49325;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE S-layer protein (surface layer protein) (Fragment).  
OS Bacillus thuringiensis (subsp. galleriae).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=29338;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=NRRL 4045; PubMed=2592346;  
RX MEDLINE=90078111; PubMed=2592346;  
RA Luckeich M.D.; Beveridge T.J.;  
RT "Characterization of a dynamic S layer on Bacillus thuringiensis.";  
RL J. Bacteriol. 171:6656-6667(1989).  
CC -I- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY  
CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.  
CC -I- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A  
CC S-LAYER WITH OBLIQUE (P2) SYMMETRY.  
CC Cell wall; S-layer.  
KW NON TER 10 10  
SQ SEQUENCE 10 AA; 1080 MW; 57AECACAB769D1A3 CRC64;

Query Match 23.5%; Score 19; DB 1; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 GKTY 14  
: |||:  
Db 2 GKTF 5

RESULT 7  
SAMP MUSCA  
ID -SAMP MUSCA STANDARD; PRT; 9 AA.  
AC P19055;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Serum amyloid P-component (SAP) (Fragment).  
OS Mustelus canis (Smooth dogfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;  
OC Mustelus.  
OX NCBI\_TaxID=7812;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=83160932; PubMed=6403520;  
RA Robey F.A.; Tanaka T.; Liu T.-Y.;  
RT "Isolation and characterization of two major serum proteins from the  
RT dogfish, Mustelus canis, C-reactive protein and amyloid P  
RT component.";  
RL J. Biol. Chem. 258:3889-3894(1983).  
CC -I- SUBUNIT: HOMOPENTAMER. PENTAXIN (OR PENTRAXIN) HAVE A DISCOID  
CC ARRANGEMENT OF 5 NONCOVALENTLY BOUND SUBUNITS.  
CC -I- DISEASE: SAP IS A PRECURSOR OF AMYLOID COMPONENT P WHICH IS FOUND  
CC IN BASEMENT MEMBRANE AND ASSOCIATED WITH AMYLOID DEPOSITS.  
CC -I- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.

DR PIR; B20569; B20569.  
DR InterPro; IPR001759; Pentaxin.  
DR PROSITE; PS00289; PENTAXIN; PARTIAL.  
KW Lectin; Amyloid; Glycoprotein; Plasma; Pentaxin.  
FT DOMAIN 1 >9 PENTAXIN.  
FT NON TER 9 9  
SQ SEQUENCE 9 AA; 965 MW; D05B5735B3386769 CRC64;

Query Match 22.2%; Score 18; DB 1; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.1e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 KSLLY 8  
: |||:  
Db 5 KSLIF 9

RESULT 8  
PPCK\_FASHE  
ID -PPCK\_FASHE STANDARD; PRT; 10 AA.  
AC P80525;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Putative phosphoenolpyruvate carboxylase [GTP] (EC 4.1.1.32)  
DE (Phosphoenolpyruvate carboxylase) (PEPCK) (Newly excysted juvenile  
DE protein 1) (Fragment).  
OS Fasciola hepatica (Liver fluke).  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Fasciola.  
OC Echinostomida; Echinostomata; Fascioloides; Fasciolidae;  
OX NCBI\_TaxID=6192;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95366993; PubMed=7639732;  
RA Tkalcovic J.; Ashman K.; Meeusen E.;  
RT "Fasciola hepatica: rapid identification of newly excysted juvenile  
RT proteins.";  
RL Biochem. Biophys. Res. Commun. 213:169-174(1995).  
CC -I- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate  
CC + CO(2).  
CC -I- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]  
CC FAMILY.  
DR InterPro; IPR000364; PEP\_carboxykin.  
DR PROSITE; PS00505; PEPCK GTP; PARTIAL.  
KW Lyase; Decarboxylase; GTP-binding.  
FT NON TER 10 10  
SQ SEQUENCE 10 AA; 1069 MW; 8393A6187AA9C87A CRC64;

Query Match 22.2%; Score 18; DB 1; Length 10;  
Best Local Similarity 75.0%; Pred. No. 2.1e+03;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 YKDG 11  
: |||:  
Db 1 YPDG 4

RESULT 9  
SPI\_HALRO  
ID -SPI\_HALRO STANDARD; PRT; 10 AA.  
AC Q10597;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Serine proteinase inhibitor (Fragment).  
OS Halocynthia roretzi (Sea squirt).  
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
OC Stolidobranchia; Pyridae; Halocynthia.  
OX NCBI\_TaxID=7729;  
RN [1]  
RP SEQUENCE.  
RX TISSUE=Hemolymph;  
RX MEDLINE=96321313; PubMed=8759295;

```

RA Shishikura F., Abe T., Ohtake S.-I., Tanaka K.;
RT "Purification and characterization of a 58,000-Da proteinase
RT inhibitor from the hemolymph of a solitary ascidian, Halocynthia
RL roretzi.";
RL Comp. Biochem. Physiol. 114B:1-9(1996).
CC -1- FUNCTION: STRONGLY INHIBITS TRYPSIN AND PLASMA ENZYME(S) ACTIVITY.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR InterPro: IPR000215; Serpin.
DR PROSITE: PS00284; SERPIN; PARTIAL.
KW Serpin; Serine protease inhibitor; glycoprotein; Plasma.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1104 MW; 4225C73B1B187AA3 CRC64;

Query Match 22.2%; Score 18; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 2.1e+03; Indels 0; Gaps 0;
Matches 3; Conservative 1; Mismatches 0;

QY 9 KQCK 12
DB 3 KDGE 6

RESULT 10
CXAL_CONMA STANDARD; PRT; 14 AA.
AC P01521;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-conotoxin MI (MI).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6492;
RN [1]
RP SEQUENCE.
RA McLintosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
RX "Isolation and structure of a peptide toxin from the marine snail
RT Conus magus.";
RL Arch. Biochem. Biophys. 218:329-334(1982).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84032400; PubMed=6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin MI. Disulfide bonding and conformational states.";
RL J. Biol. Chem. 258:12247-12251(1983).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON/POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC FAMILY.
DR PIR; A01784; NTKNIM.
DR HSSP; P56973; I845.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation.
FT DISULFID 3 8
FT DISULFID 4 14
FT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1499 MW; DEE91898B8F5E5BD CRC64;

Query Match 22.2%; Score 18; DB 1; Length 14;
Best Local Similarity 75.0%; Pred. No. 3e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 GKTY 14
DB 9 GKTY 12

```

```

RESULT 11
EFTU_CANPA STANDARD; PRT; 14 AA.
AC P54835;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Elongation factor Tu, mitochondrial (Fragment).
GN TUFM.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
DR HSC-2DPAGE; P54835; DOG.
DR InterPro: IPR000795; EF_GTPbind.
DR PROSITE: PS00301; EFACFOR_GTP; PARTIAL.
KW Elongation factor; Protein biosynthesis; Mitochondrion; GTP-binding.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1600 MW; 8CAF0B6AE7CDE41 CRC64;

Query Match 22.2%; Score 18; DB 1; Length 14;
Best Local Similarity 75.0%; Pred. No. 3e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 KTYL 15
DB 6 KTYV 9

RESULT 12
MCRZ_METTM STANDARD; PRT; 14 AA.
AC P58816;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methyl-coenzyme M reductase II gamma subunit (EC 1.8.-.-) (MCR II
DE gamma) (Fragment).
GN MRTG.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=79929;
RN [1]
RP SEQUENCE.
RX MEDLINE=91099370; PubMed=2269306;
RA Rospert S., Linder D., Ellermann J., Thauer R.K.;
RT "Two genetically distinct methyl-coenzyme M reductases in
RT Methanobacterium thermoautotrophicum strain Marburg and delta H.";
RL Eur. J. Biochem. 194:871-877(1990).
CC -1- FUNCTION: Reduction of methyl-coenzyme M (2-(methylthio)
CC ethanesulfonic acid) with 7-mercaptoheptanoylthreonine phosphate
CC to methane and an heterodisulfide.
CC -1- CATALYTIC ACTIVITY: CH(3)-S-CoM + H-S-HTP = CH(4) + CoM-S-S-HTP.
CC -1- COFACTOR: THE ENZYME COMPLEX BINDS TIGHTLY (BUT NOT COVALENTLY)
CC TO TWO MOLECULES OF COENZYME F430. F430 IS A YELLOW NICKEL
CC PORPHINOID.
CC -1- PATHWAY: Methanogenesis; last step.
CC -1- SUBUNIT: HEXAMER OF TWO ALPHA, TWO BETA, AND TWO GAMMA CHAINS.

```



CC -1- DEVELOPMENTAL STAGE: THERE ARE TWO MCR COMPLEXES IN THIS BACTERIA.  
 CC MCR II IS EXPRESSED IN THE EARLY GROWTH PHASE. LATE GROWTH CELLS  
 CC CONTAINS MOSTLY MCR I.  
 KW Methanogenesis; Oxidoreductase; Multigene family.  
 FT INIT MET 0  
 FT NON TER 14  
 SQ SEQUENCE 14 AA; 1557 MW; 97E9439C4223B871 CRC64;  
 Query Match 22.2%; Score 18; DB 1; Length 14;  
 Best Local Similarity 37.5%; Pred. No. 3e+03;  
 Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 8 YKDGKTYL 15  
 DB 6 YTPGETQI 13  
 RESULT 13  
 NIS1\_ANAVA STANDARD; PRT; 15 AA.  
 AC Q44507;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cysteine desulfurase 1 (EC 4.4.1.-) (Nitrogenase metalloclusters  
 DE biosynthesis protein nifS1) (Fragment).  
 GN NIFS1 OR NIFS.  
 OS Anabaena variabilis.  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 OX NCBI\_TaxID=1172;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ECC 7937 / ATCC 29413;  
 RA Monnerjahn U.; Boehme H.;  
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP EXPRESSION PATTERN.  
 RC STRAIN=PCC 7937 / ATCC 29413;  
 RX MEDLINE=96016168; PubMed=7568132;  
 RA Thiel T.; Lyons E.M.; Erker J.C.; Ernst A.;  
 RT "A second nitrogenase in vegetative cells of a heterocyst-forming  
 cyanobacterium."  
 CC -1- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR ATOMS FROM  
 CYSTEINE TO PRODUCE ALANINE. SEEMS TO PARTICIPATE IN THE  
 BIOSYNTHESIS OF THE NITROGENASE METALLOCLUSTERS BY PROVIDING THE  
 INORGANIC SULFUR REQUIRED FOR THE FE-S CORE FORMATION (BY  
 SIMILARITY).  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: BELONGS TO THE NIF1 GENE CLUSTER WHICH IS EXPRESSED  
 IN HETEROCYST UNDER ANAEROBIC AND AEROBIC CONDITIONS.  
 CC -1- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 AMINOTRANSFERASES. NIFS/ISCS SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X69898; CAA49523.1; -;  
 DR InterPro: IPR000192; AminoTransf.  
 DR PROSITE; PS00595; AA\_TRANSFER CLASS\_5; PARTIAL.  
 KW Nitrogen fixation; Lyase; Pyridoxal phosphate.  
 FT NON TER 15  
 SQ SEQUENCE 15 AA; 1684 MW; 08B8F106DE65547D CRC64;  
 Query Match 22.2%; Score 18; DB 1; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 3.3e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 SLLYKD 10  
 DB 2 SVIYLD 7  
 RESULT 14  
 GON1\_PETWA STANDARD; PRT; 10 AA.  
 AC P04378;  
 DT 20-MAR-1987 (Rel. 04, Created)  
 DT 20-MAR-1987 (Rel. 04, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I)  
 DE (Luliberin I).  
 OS Petromyzon marinus (Sea lamprey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
 OX NCBI\_TaxID=7757;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=86168192; PubMed=3514603;  
 RA Sherwood N.M.; Sower S.A.; Marshak D.R.; Fraser B.A.; Brownstein M.J.;  
 RT "Primary structure of gonadotropin-releasing hormone from lamprey  
 brain."  
 RL J. Biol. Chem. 261:4812-4819(1986).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 FOLLICULAR-STIMULATING HORMONES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 DR PIR; A01412; RHLMSG.  
 DR InterPro: IPR002012; GNRH.  
 DR Pfam; PF00446; GNRH; 1.  
 DR PROSITE; PS00473; GNRH; 1.  
 KW Hormone; Amidation; Hypochalamus.  
 FT MOD RES 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1244 MW; 1E4B36237B1735AB CRC64;  
 Query Match 21.0%; Score 17; DB 1; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 5 SLLYKDG 11  
 DB 4 SLEWKP 10  
 RESULT 15  
 SP34\_DICMU STANDARD; PRT; 10 AA.  
 AC P81545;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Surface protein P34 (Fragment).  
 GN P34.  
 OS Dictyostelium mucoroides (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
 OX NCBI\_TaxID=31287;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=DM-7;  
 RA Schreiner S.J.;  
 RT "Characterization of a surface protein in macrocysts of Dictyostelium  
 mucoroides."  
 RL Submitted (NOV-1998) to the SWISS-PROT data bank.  
 CC -1- FUNCTION: THIS PROTEIN IS PRESENT IN THE MACROCYST PRIMARY WALL  
 WHICH IS PRODUCED BY AMOEBAE DURING THE ONSET OF SEXUAL  
 REPRODUCTION.  
 CC -1- SUBCELLULAR LOCATION: MACROCYST PRIMARY WALL.  
 KW Cell wall.

FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1190 MW; 1B6A707AA3345B50 CRC64;  
Query Match 21.0%; Score 17; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 9 KDG 11  
Db 6 KDG 8

Search completed: April 22, 2003, 13:13:03  
Job time : 25 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:10:19 ; Search time 28 Seconds  
(without alignments)  
117.741 Million cell updates/sec

Title: US-09-674-716B-3

Perfect score: 81  
Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 3565

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP\_TREMBL 21:\*
- 2: sp\_archaea:\*
- 3: sp\_bacteria:\*
- 4: sp\_fungi:\*
- 5: sp\_human:\*
- 6: sp\_invertebrate:\*
- 7: sp\_mammal:\*
- 8: sp\_mhc:\*
- 9: sp\_organelle:\*
- 10: sp\_phage:\*
- 11: sp\_plant:\*
- 12: sp\_rodent:\*
- 13: sp\_virus:\*
- 14: sp\_vertebrate:\*
- 15: sp\_unclassified:\*
- 16: sp\_rvirus:\*
- 17: sp\_bacteriaph:\*
- 18: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	33.3	12	11 Q61331	Q61331 mus musculus
2	27	33.3	15	5 P82207	P82207 bombyx mori
3	23	28.4	8	4 Q9UJ50	Q9UJ50 homo sapien
4	23	28.4	16	2 Q9X7K1	Q9X7K1 rhodobacter
5	22	27.2	12	11 Q91YF5	Q91YF5 mus musculus
6	22	27.2	13	4 Q16141	Q16141 homo sapien
7	22	27.2	15	2 Q9R5L9	Q9R5L9 bacillus st
8	22	27.2	15	5 Q9TW6	Q9TW6 rapana thom
9	22	27.2	16	4 Q16183	Q16183 homo sapien
10	21	25.9	12	8 P82164	P82164 spinacia ol
11	21	25.9	15	2 P82468	P82468 pseudomonas
12	21	25.9	15	7 Q9TNP2	Q9TNP2 homo sapien
13	20	24.7	11	2 Q9L4F7	Q9L4F7 bacillus ce
14	20	24.7	14	5 Q9NFK8	Q9NFK8 brugia paha
15	20	24.7	15	10 Q9S8R5	Q9S8R5 gosypium h
16	20	24.7	16	4 Q9UD47	Q9UD47 homo sapien

17	20	24.7	16	4 Q9UC18	Q9UC18 homo sapien
18	19	23.5	9	4 Q9UKJ6	Q9UKJ6 homo sapien
19	19	23.5	11	5 P83321	P83321 penaeus mon
20	19	23.5	11	13 Q90WA2	Q90WA2 gallus gall
21	19	23.5	13	16 Q9K9D7	Q9K9D7 bacillus ha
22	19	23.5	14	10 Q945F2	Q945F2 cicier ariet
23	19	23.5	14	10 Q8S2V1	Q8S2V1 zea mays (m
24	19	23.5	15	3 Q9URE0	Q9URE0 saccharomyc
25	19	23.5	15	6 Q9TFT5	Q9TFT5 sus scrofa
26	19	23.5	15	6 Q9TFT3	Q9TFT3 sus scrofa
27	19	23.5	15	10 Q8S2V2	Q8S2V2 zea mays (m
28	18	22.2	11	5 Q95FX6	Q95FX6 caenorhabdi
29	18	22.2	12	10 Q38715	Q38715 arachis hyp
30	18	22.2	13	6 Q9GJU2	Q9GJU2 ovis aries
31	18	22.2	13	13 Q8QGZ5	Q8QGZ5 fugu rubrip
32	18	22.2	14	10 P83291	P83291 arabidopsis
33	18	22.2	15	5 Q9TWS0	Q9TWS0 pyura stolo
34	18	22.2	15	12 Q91PF4	Q91PF4 rabies viru
35	18	22.2	15	12 Q91PES	Q91PES rabies viru
36	18	22.2	15	12 Q91PD8	Q91PD8 rabies viru
37	18	22.2	16	4 Q96KQ2	Q96KQ2 homo sapien
38	18	22.2	16	8 Q35214	Q35214 oenothera b
39	17	21.0	8	6 Q9BFA7	Q9BFA7 macruscelid
40	17	21.0	10	2 Q44693	Q44693 bacillus am
41	17	21.0	10	6 Q9TRC1	Q9TRC1 bos taurus
42	17	21.0	11	2 Q9AIY6	Q9AIY6 carsonella
43	17	21.0	12	6 Q9TRT8	Q9TRT8 bos taurus
44	17	21.0	12	8 P92457	P92457 ephedra sp.
45	17	21.0	12	10 Q02320	Q02320 pinus sylve

## ALIGNMENTS

RESULT 1  
ID Q61331 PRELIMINARY; PRT; 12 AA.  
AC Q61331;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
DE N-acetylglucosamine (Betal-4) galactosyl transferase (EC 2.4.1.90)  
DE (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89033997; PubMed=3141392;  
RA Nakazawa K., Ando T., Kimura T., Narimatsu H.;  
RT "Cloning and sequencing of a full-length cDNA of mouse N-  
acetylglucosamine (betal-4)galactosyltransferase.";  
RL J. Biochem. 104:165-168(1988).  
DR EMBL; D00315; BAA00217.1; -;  
KW Glycosyltransferase; Transferase.  
FT NON\_TER 1  
FT NON\_TER 12  
SQ SEQUENCE 12 AA; 1283 MW; 304EA40668387728 CRC64;

Query Match 33.3%; Score 27; DB 11; Length 12;  
Best Local Similarity 100.0%; Pred.No. 5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 GKTLYL 15  
|||  
Db 7 GKTLYL 11

RESULT 2  
ID P82207 PRELIMINARY; PRT; 15 AA.  
AC P82207;

DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE Unknown protein from 2D-page (Fragment).  
 OS Bombyx mori (silk moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Bombycoidea; Bombycidae; Bombyx.  
 OX NCBI\_TaxID=7091;  
 RN [1]  
 RN SEQUENCE.  
 RP STRAIN=XINHANG X KEMING; TISSUE=BODY WALL, AND FAT BODY;  
 RX MEDLINE=21177481; PubMed=11280994;  
 RA Zhong B.X.;  
 RT "Protein database for several tissues derived from five instar of  
 RT silkworm".  
 RL i Chuan Hsueh Pao 28:217-224 (2001).  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1877 MW; 590F6BD4703CA70C CRC64;  
 Query Match 33.3%; Score 27f DB 5; Length 15;  
 Best Local Similarity 57.1%; Pred. No. 6.5e+02;  
 Matches 8; Conservative 1; Mismatches 3; Indels 2; Gaps 1;  
 QY 3 SKSLYKDGKTYLN 16  
 DB 1 SKSLFYQ--KQYDN 12  
 RESULT 3  
 Q9UJ50 PRELIMINARY; PRT; 8 AA.  
 ID Q9UJ50;  
 AC Q9UJ50;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Latrophilin-2 (Fragment).  
 GN LPFH1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99153747; PubMed=10030676;  
 RX White G.R.M., Varley J.M., Heighway J.;  
 RA "Isolation and characterisation of a human homologue of the  
 RT latrophilin gene from a region of lp31.1 implicated in breast  
 RT cancer.";  
 RL Oncogene 17:3513-3519 (1998).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=20225451; PubMed=10760572;  
 RX White G.R.M., Varley J.M., Heighway J.;  
 RA "Genomic structure and expression profile of LPFH1, a 7TM gene  
 RT variably expressed in breast cancer cell lines.";  
 RL Biochim. Biophys. Acta 1491:75-92 (2000).  
 DR EMBL; AJ244509; CAB60204.1; --  
 FT NON\_TER 1 1  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 1025 MW; 4D59D763344AAB0B CRC64;  
 Query Match 28.4%; Score 23; DB 4; Length 8;  
 Best Local Similarity 37.5%; Pred. No. 6.7e+05;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 8 YKDGKTYL 15  
 DB 1 YEDNKPF 8  
 RESULT 4  
 Q9X7K1 PRELIMINARY; PRT; 16 AA.  
 ID Q9X7K1;  
 AC Q9X7K1;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DE Molybdopterin cofactor biosynthesis protein C C-terminus  
 DE (Fragment).  
 GN MOBC.  
 OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
 OC Rhodobacter.  
 OX NCBI\_TaxID=1061;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=B10S;  
 RC Lemkueller S., Klipp W.;  
 RT "The molybdenum cofactor biosynthesis protein MoBA from Rhodobacter  
 RT capsulatus is required for the activity of molybdenum enzymes binding  
 RT MGD, but not for xanthine dehydrogenase harbouring the MPT cofactor.";  
 RL FEBS Lett. 174:239-246 (1999).  
 DR EMBL; AJ1131528; CAB43542.1; --  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 SQ SEQUENCE 16 AA; 1692 MW; 1DFE0534394788F8 CRC64;  
 Query Match 28.4%; Score 23; DB 2; Length 16;  
 Best Local Similarity 83.3%; Pred. No. 3.2e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 6 LLYKDG 11  
 DB 3 LLSKDG 8  
 RESULT 5  
 Q91YF5 PRELIMINARY; PRT; 12 AA.  
 ID Q91YF5;  
 AC Q91YF5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE GM3 synthase protein (Fragment).  
 GN GM3 SYNTHASE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=ICR; TISSUE=BRAIN;  
 RC Shuichi T.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y18020; CAC79652.1; --  
 FT NON\_TER 12 12  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1445 MW; 84E7876609572735 CRC64;  
 Query Match 27.2%; Score 22; DB 11; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 3.5e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 SLLYKD 10  
 DB 5 SULLKD 10  
 RESULT 6  
 Q16141 PRELIMINARY; PRT; 13 AA.  
 ID Q16141;  
 AC Q16141;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Mutant DNA polymerase beta (Fragment).  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94220089; PubMed=7545922;  
RA Sadakane Y., Maeda K., Kuroda Y., Hori K.;  
RT "Identification of mutations in DNA polymerase beta mRNAs from  
RT patients with Werner syndrome";  
RL Biochem. Biophys. Res. Commun. 200:219-225(1994).  
DR EMBL; S69873; AAD14051.1; --  
FT NON\_TER 1  
SQ SEQUENCE 13 AA; 1568 MW; D88C62798C9542CD CRC64;  
  
Query Match 27.2%; Score 22; DB 4; Length 13;  
Best Local Similarity 37.5%; Pred. No. 3.8e+03;  
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 9 KDGKTYLN 16  
Db :||:|:  
3 RDSAVVIN 10  
  
RESULT 7  
ID Q9SL9 PRELIMINARY; PRT; 15 AA.  
AC Q9SL9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE EXO-alpha-1,4-glucosidase [EC 3.2.1.20] (Fragment).  
OS Bacillus stearothermophilus.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Geobacillus.  
OX NCBI\_TaxID=1422;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92209510; PubMed=1555585;  
RA Suzuki Y., Yonezawa K., Hattori M., Takii Y.;  
RT "Assignment of Bacillus thermoanaerobius KPI071 alpha-  
RT glucosidase I to an exo-alpha-1,4-glucosidase, and its striking  
RT similarity to bacillary oligo-1,6-glucosidases in N-terminal sequence  
RT and in structural parameters calculated from the amino acid  
RT composition";  
RL Eur. J. Biochem. 205:249-256(1992).  
SQ SEQUENCE 15 AA; 1931 MW; 62B4CE40013D3042 CRC64;  
  
Query Match 27.2%; Score 22; DB 2; Length 15;  
Best Local Similarity 36.4%; Pred. No. 4.4e+03;  
Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
  
QY 4 KSLLYKDGKTY 14  
Db :||:|:  
2 KKTWWKEGVAY 12  
  
RESULT 8  
ID Q9TWR6 PRELIMINARY; PRT; 15 AA.  
AC Q9TWR6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE Hemocyanin RHSS1 subunit (Fragment).  
OS Rapana thomasi (Marine snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Muricoidea; Muricidae; Rapana.  
OX NCBI\_TaxID=29165;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94007762; PubMed=8403854;  
RA Idakieva K., Severov S., Svendsen I., Genov N., Stoeva S.,  
RA Beltramini M., Tognon G., Di Muro P., Salvato B.;

RT "Structural properties of Rapana thomasi grosse hemocyanin:  
RT isolation, characterization and N-terminal amino acid sequence of two  
RT different dissociation products."; (1993).  
RL Comp. Biochem. Physiol. 106B:53-59 (1993).  
FT NON\_TER 1  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 1735 MW; 78985413C9E90B6B CRC64;  
  
Query Match 27.2%; Score 22; DB 5; Length 15;  
Best Local Similarity 55.6%; Pred. No. 4.4e+03;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 SLLYKDGKT 13  
Db :||:|:  
1 SLPYKRVRT 9  
  
RESULT 9  
ID Q16183 PRELIMINARY; PRT; 16 AA.  
AC Q16183;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE GPIa protein (Fragment).  
GN GPIa.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94378189; PubMed=7916494;  
RA Kaib R., Santoso S., Unkelbach K., Kiesel V., Mueller-Eckhardt C.;  
RT "Localization of the Br polymorphism on a 144 bp exon of the GPIa gene  
RT and its application in platelet DNA typing.";  
RL Thromb. Haemost. 71:651-654(1994).  
DR EMBL; S72155; AAD14096.1; --  
FT NON\_TER 1  
SQ SEQUENCE 16 AA; 1968 MW; B31EPE05E038C26A CRC64;  
  
Query Match 27.2%; Score 22; DB 4; Length 16;  
Best Local Similarity 42.9%; Pred. No. 4.8e+03;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 9 KDGKTYLN 15  
Db :||:|:  
5 EGRVYL 11  
  
RESULT 10  
ID P82164 PRELIMINARY; PRT; 12 AA.  
AC P82164;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Chloroplast 30S ribosomal protein S14 beta (Fragment).  
OS Spinacia oleracea (Spinach).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
RC STRAIN=CV. ALMARO; TISSUE=LEAF;  
RX MEDLINE=20435797; PubMed=10874039;  
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
RT the small subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 274:28455-28465(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.

CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -1- MISCELLANEOUS: S14 ALPHA AND BETA FORMS DIFFER IN PL. S14 ALPHA  
 CC FORM IS THE LEAST BASIC FORM.  
 CC -1- MASS SPECTROMETRY: MW=11745.9; METHOD=ELECTROSPRAY.  
 CC -1- MASS SPECTROMETRY: MW=11947; METHOD=MALDI.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 14 KDA.  
 CC -1- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.  
 DR InterPro: IPR001209; Ribosomal\_S14.  
 DR InterPro: IPR001648; Ribosomal\_S18.  
 DR Pfam: PF00253; Ribosomal\_S14; PARTIAL.  
 DR PROSITE: PS00057; RIBOSOMAL\_S18; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1513 MW; 5E84D3F10116D057 CRC64;  
 Query Match 25.9%; Score 21; DB 8; Length 12;  
 Best Local Similarity 44.4%; Pred. No. 5.1e+03;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 KSLLYKDGK 12  
 DB 3 KSLIQREKK 11  
 RESULT 11  
 ID P82468 PRELIMINARY; PRT; 15 AA.  
 AC P82468;  
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE Lysyl endopeptidase (EC 3.4.21.50) (Fragment).  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE  
 RA Lehman J.;  
 RL Submitted (APR-2000) to the SWISS-PROT data bank.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: LYS-|-XAA, INCLUDING  
 CC LYS-|-PRO.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S5 (SERINE PROTEASE).  
 KW Hydrolase; Serine protease.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1483 MW; 3F42C5E66C4F76DD CRC64;  
 Query Match 25.9%; Score 21; DB 2; Length 15;  
 Best Local Similarity 75.0%; Pred. No. 6.5e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 YKDG 11  
 DB 3 YRDG 6  
 RESULT 12  
 ID Q9TNP2 PRELIMINARY; PRT; 15 AA.  
 AC Q9TNP2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Fraction 60-HLA BC locus class I-bound peptide (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE  
 RA Harris P.E., Colovai A.I., Maffei A., Liu Z., Foca N.S.;  
 RX MEDLINE=96165039; PubMed=8567028;  
 DR EMBL; AJ277990; CAB93515.1; -.

RT "Major histocompatibility complex class I presentation of exogenous  
 RT and endogenous protein-derived peptides by a transfected human  
 RT monocyte cell line.";  
 RL Immunology 86:606-611(1995).  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1838 MW; 98DCDC01F8F7E15 CRC64;  
 Query Match 25.9%; Score 21; DB 7; Length 15;  
 Best Local Similarity 44.4%; Pred. No. 6.5e+03;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 8 YKDGKTYLN 16  
 DB 3 YGVARTYLD 11  
 RESULT 13  
 ID Q9L4F7 PRELIMINARY; PRT; 11 AA.  
 AC Q9L4F7;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE Phosphatidylinositol-specific phospholipase C (PI-PLC)  
 DE (Fragment).  
 GN PLCA.  
 OS Bacillus cereus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1396;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 14579 TYPE STRAIN;  
 RX MEDLINE=20055637; PubMed=10589720;  
 RA Oktad O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.;  
 RT "Sequence analysis of three Bacillus cereus loci under PicR-regulated  
 RT genes encoding degradative enzymes and enterotoxin.";  
 RL Microbiology 145:3129-3138(1999).  
 DR EMBL; AJ243711; CAB69804.1; -.  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1335 MW; 4277A30E20572333 CRC64;  
 Query Match 24.7%; Score 20; DB 2; Length 11;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 SSKSLLYK 9  
 DB 2 SNKKLILK 9  
 RESULT 14  
 ID Q9NFK8 PRELIMINARY; PRT; 14 AA.  
 AC Q9NFK8;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE MmcI protein (Fragment).  
 GN MmcI.  
 OS Brugia pahangi.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Brugia.  
 OX NCBI\_TaxID=6280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ems R.D., Thompson F., Devaney E.;  
 RT "A novel mRNA up-regulated in mammalian-derived microfilaria of  
 RT Brugia.";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ277990; CAB93515.1; -.

```
FT NON_TER 1 1
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1602 MW; 5C4C62C55AB89397 CRC64;

Query Match 24.7%; Score 20; DB 5; Length 14;
Best Local Similarity 45.5%; Pred. No. 8.8e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSKSLLYKDGK 12
   |:|:|
Db 3 SATSLMMQMK 13

RESULT 15
Q9S8R5 PRELIMINARY; PRT; 15 AA.
AC Q9S8R5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Fiber annexin (Fragment).
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Malvales; Malvaceae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE.
RX MEDLINE=94004990; PubMed=8401609;
RA Andrawis A., Solomon M., Delmer D.P.;
RT "Cotton fiber annexins: a potential role in the regulation of callose
synthase.";
RL Plant J. 3:763-772(1993).
SQ SEQUENCE 15 AA; 1689 MW; 1087950BAC4F69F9 CRC64;

Query Match 24.7%; Score 20; DB 10; Length 15;
Best Local Similarity 50.0%; Pred. No. 9.5e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSKSLLY 8
   |:|:|
Db 5 RSANQLLH 12

Search completed: April 22, 2003, 13:14:00
Job time : 30 secs
```





GenCore version 5.1.4 p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:54:59 ; Search time 73 Seconds  
(without alignments)  
29.206 Million cell updates/sec

Title: US-09-674-716B-3  
Perfect score: 81  
Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 255767

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	16	AAV32254	Light chain CDR L1
2	77	95.1	16	AAV39818	Light chain CDR1 o
3	75	92.6	16	AAW39815	Light chain CDR1 o
4	75	92.6	16	AAW39824	Light chain CDR1 o
5	69	85.2	16	AAW39875	Light chain CDR1 o
6	69	85.2	16	AAW39839	Light chain CDR1 o
7	62.5	77.2	15	AAW39821	Light chain CDR1 o
8	61	75.3	16	AAU70328	Human Kappa II lig
9	60	74.1	16	AAV14404	Peptide CDR-L1 der
10	59	72.8	16	AAB86292	Murine derived ant

11	59	72.8	16	22	AAB73652	Murine monoclonal
12	57	70.4	16	13	AAW24704	Sequence of an ant
13	55	67.9	16	14	AAW30448	C242-11 MAb kappa
14	55	67.9	16	22	AAU07734	Human light chain
15	55	67.9	16	23	AAO14984	Murine mAb 1D9 anti
16	53	65.4	16	18	AAW31747	CDRL1 region of L
17	53	65.4	16	19	AAW71894	Anti-human Fas mon
18	53	65.4	16	21	AAW12905	Anti-human Fas ant
19	53	65.4	16	21	AAW80281	Humanised anti-Fas
20	52	64.2	16	21	AAW92169	Murine 13H10 light
21	51	63.0	16	19	AAW70926	CDR1 of the light
22	49	60.5	16	16	AAW75487	Mouse antibody var
23	49	60.5	16	21	AAW21364	Mouse antibody 13G
24	49	60.5	16	22	AAW97198	Murine anti-PI-3,4
25	48	59.3	16	23	AAU70349	Mouse Kappa II lig
26	47	58.0	16	19	AAW50213	Sequence from the
27	46	56.8	16	16	AAW70451	VL sequence of ant
28	46	56.8	16	19	AAW58534	CDR-1 of the L cha
29	46	56.8	16	21	AAW51154	Murine CD4/CD34 re
30	45	55.6	16	19	AAW70920	CDR1 of the light
31	45	55.6	16	21	AAW95222	Anti-platelet glyc
32	45	55.6	16	22	AAW50790	Murine antibody S2
33	43	53.1	16	15	AAW59420	6D9 antibody light
34	43	53.1	16	22	AAW64460	Mouse 6D9 catalyti
35	41	50.6	16	22	AAW64463	L27ey catalytic an
36	40	49.4	16	14	AAW40216	Sequence of light
37	40	49.4	16	22	AAW67670	Complementarity de
38	39	48.1	16	22	AAW64461	3H5 catalytic ant
39	39	48.1	16	23	AAU11180	Mouse antibody Act
40	36	44.4	13	22	AAB82174	H-2Kd MHC-Class I
41	34	42.0	16	22	AAW64462	5H3 catalytic ant
42	32.5	40.1	15	16	AAW70195	MAB 3B9 light chai
43	32.5	40.1	15	19	AAW83027	Anti-Fas MAB HFET7A
44	32.5	40.1	15	20	AAW23772	CDR of the light c
45	32.5	40.1	15	20	AAV18114	Light chain CDR fo

ALIGNMENTS

RESULT 1  
AAV32254  
ID AAV32254 standard; Peptide; 16 AA.  
XX  
AC AAV32254;  
XX  
DT 15-FEB-2000 (first entry)  
XX  
DE Light chain CDR L1 of mouse anti-CD23 MAb C11.  
XX  
KW CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse;  
KW monoclonal antibody; chimeric antibody; humanised antibody;  
KW complementarity determining region; CDR; autoimmune disease;  
KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
KW urticaria; nephrotic syndrome; glomerulonephritis;  
KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinis;  
KW graft-versus-host disease; COPD; bronchitis; diabetes;  
KW B-cell malignancy; therapy.  
XX  
OS Mus musculus.  
XX  
PN WO9958679-A1.  
XX  
PD 18-NOV-1999.  
XX  
PF 07-MAY-1999; 99WO-GB01434.  
XX  
PR 09-MAY-1998; 98CB-0009839.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX

PI Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 XX WPI; 2000-053101/04.  
 DR N-PSDB; AA234739.  
 XX  
 PT Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 XX  
 PS Claim 1; Page 40; 81pp; English.  
 XX  
 CC This sequence represents complementarity determining region 1  
 CC (CDR L1) of the light chain of murine anti-CD23 (FCBRII) monoclonal  
 CC antibody C11 (see also AA32262). The invention provides altered  
 CC antibodies, such as chimeric or humanised antibodies, which comprise  
 CC sufficient of the amino acid sequences of C11 light and heavy chain  
 CC CDRs (see AA32254-59) to render them capable of binding to the CD23  
 CC type II molecule expressed on haematopoietic cells. The antibodies  
 CC are used to block soluble CD23 formation for treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX  
 SQ Sequence 16 AA;  
 Query Match 100.0%; Score 81; DB 21; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 5e-07;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RSSKSLLYKDGKTYLN 16  
 |||||:|||||  
 Db 1 RSSKSLLYKDGKTYLN 16  
 |||||:|||||  
 RESULT 2  
 AAW39818  
 ID AAW39818 standard; peptide; 16 AA.  
 AC AAW39818;  
 XX  
 DT 16-JUN-1998 (first entry)  
 XX  
 DE Light chain CDR1 of catalytic antibody 6A12.  
 XX  
 KW Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9749800-A1.  
 XX  
 PD 31-DEC-1997.  
 XX  
 PF 25-JUN-1997; 97WO-US10965.  
 XX  
 PR 25-JUN-1996; 96US-0672345.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Landry DW;  
 XX  
 DR WPI; 1998-077166/07.  
 XX  
 PT New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding

PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX  
 PS Claim 13; Page 81; 147pp; English.  
 XX  
 CC AAW39818-20 represent the sequences of the light chain complementarity  
 CC determining regions (CDRs) of the catalytic antibody 6A12, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody was  
 CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
 CC monoester transition state analogue. Antibody 6A12 has a per minute Kcat  
 CC of 0.072. The antibodies reduce the concentration of cocaine in a  
 CC subject, and are used particularly for the treatment of an overdose. They  
 CC are also used for treating addiction (by reducing the in vivo  
 CC concentration that can be achieved).  
 XX  
 SQ Sequence 16 AA;  
 Query Match 95.1%; Score 77; DB 19; Length 16;  
 Best Local Similarity 93.8%; Pred. No. 2.4e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RSSKSLLYKDGKTYLN 16  
 |||||:|||||  
 Db 1 RSSKSLLYKDGKTYLN 16  
 |||||:|||||  
 RESULT 3  
 AAW39815  
 ID AAW39815 standard; peptide; 16 AA.  
 XX  
 AC AAW39815;  
 XX  
 DT 16-JUN-1998 (first entry)  
 XX  
 DE Light chain CDR1 of catalytic antibody 3B9.  
 XX  
 KW Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9749800-A1.  
 XX  
 PD 31-DEC-1997.  
 XX  
 PF 25-JUN-1997; 97WO-US10965.  
 XX  
 PR 25-JUN-1996; 96US-0672345.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Landry DW;  
 XX  
 DR WPI; 1998-077166/07.  
 XX  
 PT New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX  
 PS Claim 11; Page 80; 147pp; English.  
 XX  
 CC AAW39815-17 represent the sequences of the light chain complementarity  
 CC determining regions (CDRs) of the catalytic antibody 3B9, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 3B9 antibody was

CC identified using TSA1, which is an immunogenic conjugate of a phosphate monoester transition state analogue. Antibody 3B9 has a per minute Kcat of 0.11. The antibodies reduce the concentration of cocaine in a subject, CC and are used particularly for the treatment of an overdose. They are also CC used for treating addiction (by reducing the in vivo concentration that CC can be achieved).

SQ Sequence 16 AA;  
Query Match 92.6%; Score 75; DB 19; Length 16;  
Best Local Similarity 87.5%; Pred. No. 5.2e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSKSLLYXDKGKTYLN 16  
|||:||||:|||||  
Db 1 RSRSLLYXDKGKTYLN 16

RESULT 4  
AAW39824  
ID AAW39824 standard; peptide; 16 AA.

XX AAW39824;  
XX  
DT 16-JUN-1998 (first entry)

XX Light chain CDR1 of catalytic antibody 12H1.

XX Variable domain; lambda light chain; catalytic antibody; degradation;  
KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
KW overdose; addiction.

XX Mus sp.  
XX WO9749800-A1.  
XX 31-DEC-1997.  
XX 25-JUN-1997; 97WO-US10965.  
XX 25-JUN-1996; 96US-0672345.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Landry DW;

XX WPI; 1998-077166/07.

XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding

XX Claim 17; Page 83; 147pp; English.

XX AAW39824-26 represent the sequences of the light chain complementarity  
CC determining regions (CDRs) of the catalytic antibody 12H1, which is able  
CC to degrade cocaine. A series of cocaine transition state analogues  
CC (TSAs) were prepared and used to immunise mice for production of  
CC hybridomas. Catalytic antibodies were identified by their capacity to  
CC release 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody was  
CC identified using TSA2, and has a per minute Kcat of 0.16. The antibodies  
CC reduce the concentration of cocaine in a subject, and are used  
CC particularly for the treatment of an overdose. They are also used for  
CC treating addiction (by reducing the in vivo concentration that can be  
CC achieved).

XX Sequence 16 AA;

Query Match 92.6%; Score 75; DB 19; Length 16;  
Best Local Similarity 87.5%; Pred. No. 5.2e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSKSLLYXDKGKTYLN 16  
|||:||||:|||||  
Db 1 RSRSLLYXDKGKTYLN 16

RESULT 5

AAW39875  
ID AAW39875 standard; peptide; 16 AA.

XX AAW39875;  
XX  
DT 16-JUN-1998 (first entry)

XX Light chain CDR1 of a catalytic antibody capable of degrading cocaine.  
XX Variable domain; lambda light chain; catalytic antibody; degradation;  
KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
KW overdose; addiction.

XX Mus sp.

XX Key Location/Qualifiers  
FH Misc-difference 4 /note= "not specified"  
FT Misc-difference 9 /note= "not specified"

XX WO9749800-A1.  
XX 31-DEC-1997.

XX 25-JUN-1997; 97WO-US10965.  
XX 25-JUN-1996; 96US-0672345.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Landry DW;

XX WPI; 1998-077166/07.

XX New catalytic antibodies able to decompose cocaine, single-chain  
PT analogues - used to treat cocaine overdose and addiction, required  
PT in far smaller doses than antibodies that antagonise cocaine by  
PT simply binding

XX Claim 10; Page 5; 147pp; English.

XX AAW39875-77 represent the sequences of the light chain complementarity  
CC determining regions (CDRs) of a catalytic antibody which is capable of  
CC degrading cocaine. A series of cocaine transition state analogues (TSAs)  
CC were prepared and used to immunise mice for production of hybridomas.  
CC Catalytic antibodies were identified by their capacity to release  
CC 3H-benzoic acid from 3H-phenyl cocaine. These antibodies were found  
CC to have CDRs of the present sequence. The antibodies reduce the  
CC concentration of cocaine in a subject, and are used particularly for the  
CC treatment of an overdose. They are also used for treating addiction (by  
CC reducing the in vivo concentration that can be achieved).

XX Sequence 16 AA;

Query Match 85.2%; Score 69; DB 19; Length 16;  
Best Local Similarity 87.5%; Pred. No. 5.4e-05;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSKSLLYXDKGKTYLN 16  
|||:||||:|||||  
Db 1 RSRSLLYXDKGKTYLN 16

RESULT 6

```

AAW39839
ID AAW39839 standard; peptide; 16 AA.
XX
AC AAW39839;
XX
DT 16-JUN-1998 (first entry)
XX
DE Light chain CDR1 of catalytic antibody 8G4E.
XX
KW Variable domain; lambda light chain; catalytic antibody; degradation;
KW cocaine; cocaine transition state analogue; TSA; benzoic acid;
KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;
KW overdose; addiction.
XX
OS Mus sp.
XX
FN WO9749800-A1.
XX
PD 31-DEC-1997.
XX
PF 25-JUN-1997; 97WO-US10965.
XX
PR 25-JUN-1996; 96US-0672345.
XX
PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI Landry DW;
XX
DR WPI; 1998-077166/07.
XX
PT New catalytic antibodies able to decompose cocaine, single-chain
PT analogues - used to treat cocaine overdose and addiction, required
PT in far smaller doses than antibodies that antagonise cocaine by
PT simply binding
XX
PS Claim 19; Page 87; 147pp; English.
XX
CC AAW39839-41 represent the sequences of the light chain complementarity
CC determining regions (CDRs) of the catalytic antibody 8G4E, which is able
CC to degrade cocaine. A series of cocaine transition state analogues
CC (TSAs) were prepared and used to immunise mice for production of
CC hybridomas. Catalytic antibodies were identified by their capacity to
CC release 3H-benzoic acid from 3H-phenyl cocaine. The 8G4E antibody was
CC identified using TSA3, and has a per minute Kcat of 0.12. The antibodies
CC reduce the concentration of cocaine in a subject, and are used
CC particularly for the treatment of an overdose. They are also used for
CC treating addiction (by reducing the in vivo concentration that can be
CC achieved).
XX
SQ Sequence 16 AA;
XX
Query Match 85.2%; Score 69; DB 19; Length 16;
Best Local Similarity 81.2%; Pred. No. 5.4e-05;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
Db 1 KSSQSLLYSDGKTYLN 16
:|||||:|||||

RESULT 7
AAW39821
ID AAW39821 standard; peptide; 15 AA.
XX
AC AAW39821;
XX
DT 16-JUN-1998 (first entry)
XX
DE Light chain CDR1 of catalytic antibody 2A10.
XX
KW Variable domain; lambda light chain; catalytic antibody; degradation;
KW cocaine; cocaine transition state analogue; TSA; benzoic acid;
KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;

```

```

KW overdose; addiction.
XX
OS Mus sp.
XX
FN WO9749800-A1.
XX
PD 31-DEC-1997.
XX
PF 25-JUN-1997; 97WO-US10965.
XX
PR 25-JUN-1996; 96US-0672345.
XX
PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI Landry DW;
XX
DR WPI; 1998-077166/07.
XX
PT New catalytic antibodies able to decompose cocaine, single-chain
PT analogues - used to treat cocaine overdose and addiction, required
PT in far smaller doses than antibodies that antagonise cocaine by
PT simply binding
XX
PS Claim 15; Page 82; 147pp; English.
XX
CC AAW39821-23 represent the sequences of the light chain complementarity
CC determining regions (CDRs) of the catalytic antibody 2A10, which is able
CC to degrade cocaine. A series of cocaine transition state analogues
CC (TSAs) were prepared and used to immunise mice for production of
CC hybridomas. Catalytic antibodies were identified by their capacity to
CC release 3H-benzoic acid from 3H-phenyl cocaine. The 2A10 antibody was
CC identified using TSA1, which is an immunogenic conjugate of a phosphate
CC monoester transition state analogue. Antibody 2A10 has a per minute Kcat
CC of 0.011. The antibodies reduce the concentration of cocaine in a
CC subject, and are used particularly for the treatment of an overdose. They
CC are also used for treating addiction (by reducing the in vivo
CC concentration that can be achieved).
XX
SQ Sequence 15 AA;
XX
Query Match 77.2%; Score 62.5; DB 19; Length 15;
Best Local Similarity 87.5%; Pred. No. 0.00064;
Matches 14; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 RSSKSLLYKDGKTYLN 16
Db 1 RSSKS-LYEDGKTYLN 15
:|||||:|||||

RESULT 8
AAU70328
ID AAU70328 standard; Peptide; 16 AA.
XX
AC AAU70328;
XX
DT 14-FEB-2002 (first entry)
XX
DE Human Kappa II light chain CDR1.
XX
KW Immunoglobulin; antibody; light chain; heavy chain; CDR; FR;
KW complementarity determining region; framework region; IgBP;
KW transgenic plant; immunoglobulin binding protein array;
KW IgM; IgG; IgA; IgD; IgE; IgY; IgW; kappa; lambda; CHBP.
XX
OS Homo sapiens.
XX
FN WO200183806-A1.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US14349.
XX
PR 02-MAY-2000; 2000US-0563222.

```

XX (EPIC-) EPICYTE PHARM INC.  
 XX Hiatt AC, Hein MB;  
 XX WPI; 2002-055482/07.  
 XX Preparing immunoglobulin binding protein array in plant cells by  
 PT transforming the cells with different polynucleotides encoding binding  
 PT protein polypeptides specific to ligand, selecting plant cells for  
 PT preparing array -  
 XX Disclosure; Page 14; 129pp; English.  
 XX The invention relates to transforming a population of cells (e.g. plant  
 CC cells), comprising using a library of two different polynucleotides  
 CC encoding different immunoglobulin binding protein (IgBP) polypeptides  
 CC that specifically bind to a ligand or form one or more disulphide bonds  
 CC with polypeptides in transfected cells, to generate an IgBP that  
 CC binds to a ligand, and transformed plant cells are selected, and  
 CC preparing an IgBP array in plant cells. At least one peptide sequence has  
 CC at least 75% sequence identity to a framework region (FR) of a native  
 CC IgM, IgG, IgA, IgD, IgE, IgY, kappa or lambda immunoglobulin molecule.  
 CC The method is useful for preparing an immunoglobulin binding protein  
 CC array, preferably heavy chain binding protein (CHBP) array in eukaryotic  
 CC cells, especially plant cells (e.g. Agrobacterium tumefaciens or maize)  
 CC or other eukaryotic cells (e.g. insect cells or mammalian cells). The  
 CC CHBP is useful for discovery of e.g. screening assays of IgBPs having  
 CC desired characteristics. The present sequence is a mammalian  
 CC immunoglobulin derived peptide that may be incorporated into an IgBP of  
 CC the invention.  
 XX Sequence 16 AA;  
 SQ  
 Query Match 75.3%; Score 61; DB 23; Length 16;  
 Best Local Similarity 75.0%; Pred. No. 0.0012;  
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 RSSKSLLYKDGKTYLN 16  
 DB 1 RSSQSLHLHSDGNTYL 16  
 RESULT 9  
 AAY14404  
 ID AAY14404 standard; peptide; 16 AA.  
 XX  
 AC AAY14404;  
 XX  
 XX 17-AUG-1999 (first entry)  
 XX  
 DE Peptide CDR-L1 derived from anti-HCV Ser/Thr protease MAb 8D4.  
 XX  
 KW Complementarity determining region; CDR; monoclonal antibody; MAb;  
 KW hepatitis C virus; HCV; protease; binding site.  
 XX  
 OS Synthetic.  
 XX  
 XX JP1127861-A.  
 XX  
 PD 18-MAY-1999.  
 XX  
 XX 29-OCT-1997; 97JP-0297451.  
 XX  
 XX 29-OCT-1997; 97JP-0297451.  
 XX  
 XX (NIHA) JAPAN ENERGY CORP.  
 XX  
 XX WPI; 1999-350322/30.  
 XX  
 DR Neutralized antibody partial peptide derived from hepatitis C virus  
 PT - useful for inhibiting Hepatitis C Virus (HCV) serine protease  
 PT activity

XX Example 1; Page 13; 32pp; Japanese.  
 XX  
 CC This sequence represents a peptide derived for the sequence of the light  
 CC chain variable region complementarity determining region (CDR)-1 of the  
 CC anti-hepatitis C virus (HCV) Ser/Thr protease monoclonal antibody (MAb)  
 CC 8D4 protein. The invention relates to the use of partial peptides  
 CC (AAY14348-Y14353) from the MAb 8D4 for inhibiting HCV serine protease  
 CC activity.  
 XX Sequence 16 AA;  
 SQ  
 Query Match 74.1%; Score 60; DB 20; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 0.0018;  
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 RSSKSLLYKDGKTYLN 15  
 DB 1 RSSKSLHLHSDGNTYL 15  
 RESULT 10  
 AAB86292  
 ID AAB86292 standard; peptide; 16 AA.  
 XX  
 AC AAB86292;  
 XX  
 DT\* 13-SEP-2001 (first entry)  
 XX  
 DE Murine derived antibody Wue-1 light chain variable region CDR-1.  
 XX  
 KW Antibody; Wue-1; variable region; light chain; heavy chain; antitumor;  
 KW immunomodulatory; cognate antigen identification; autoimmune disease;  
 KW tumor; multiple myeloma; lymphoma; plasmacytoma; CDR-1.  
 XX  
 OS Mus sp.  
 XX  
 XX DE19962583-A1.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 23-DEC-1999; 99DE-1062583.  
 XX  
 XX 23-DEC-1999; 99DE-1062583.  
 XX  
 XX (MUEL/) MUELLER-HERMELINK H K.  
 XX (GREI/) GREINER A.  
 XX  
 PI Mueller-Hermelink HK, Greiner A;  
 XX  
 XX WPI; 2001-426596/46.  
 XX  
 XX New antibodies specific for plasma cells; useful for treatment and  
 PT diagnosis of autoimmune diseases and plasma cell tumors -  
 XX  
 PS Claim 1; Page 9; 18pp; German.  
 XX  
 CC This invention describes novel antibodies (Ab) in which the variable  
 CC region (VR) of at least one chain and/or the VR of at least one heavy  
 CC chain includes at least one of 7 specified sequences, or fragments of  
 CC these sequences, or contain at least one light chain and/or heavy  
 CC chain encoded by specific nucleic acid sequences (I) and (II),  
 CC reproduced, or their fragments. The products of the invention have  
 CC antitumor and immunomodulatory activity. Ab, or other antibodies that  
 CC recognize the same antigen, are used: (i) to identify cognate antigens;  
 CC (ii) for specific labeling of plasma cells (PC) for generating or  
 CC separation, e.g. in an extracorporeal system; (iii) for treating autoimmune  
 CC diseases and/or tumors, e.g. multiple myeloma, lymphoma and/or  
 CC plasmacytoma. Ab are specific for mature PC, i.e. they do not recognize  
 CC precursor stages, even though these precursors are used as immunogens. As  
 CC therapeutic agents, they should show fewer side effects than conventional  
 CC chemotherapeutic agents. This sequence represents the Wue-1 antibody

CC variable region light chain complementarity determining region CDR1  
XX fragment described in the method of the invention.  
SQ Sequence 16 AA;

Query Match 72.8%; Score 59; DB 22; Length 16;  
Best Local Similarity 75.0%; Pred. No. 0.0027;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSKSKLLYKDGKTYLN 16  
:||||| |||||  
Db 1 KSSQSLSDSGKTYLN 16

RESULT 11  
AAB73652  
ID AAB73652 standard; peptide; 16 AA.  
XX AC AAB73652;  
XX DT 11-SEP-2001 (first entry)  
XX DE Murine monoclonal antibody IOR C5 light chain variable region CDR1.  
XX KW Monoclonal antibody IOR C5; hybridoma ECCC 97061101; mouse; humanised;  
KW IOR C2 antigen; colorectal cancer; tumour; metastasis; therapy;  
KW drug targeting; imaging; diagnosis; cytostatic; VL;  
KW light chain variable region;  
XX Mus sp.  
XX Synthetic.  
XX OS WO200136485-A2.  
XX FN 25-MAY-2001.  
XX PD 16-NOV-2000; 2000WO-CU000004.  
XX PF 16-NOV-1999; 99CU-0000196.  
XX PR (IMMU-) CENT IMMUNOLOGIA MOLECULAR.  
XX PA Mateo De Acosta Del Rio CM, Roque Navarro LT, Morales Morales A;  
XX PI Perez Rodriguez R, Ayala Avila M, Gavilondo Cowley JV;  
XX PI Duenas Porto M, Bell Garcia H, Rengifo Calzado E, Iznaga Escobar N;  
XX PI Ramos Zurate M;  
XX WPI; 2001-343804/36.  
XX New recombinant antibody, or single-chain fragment, derived from murine  
PT monoclonal IOR C5, useful for treatment and diagnosis of colo-rectal  
PT cancer -  
XX Claim 2; Page 24; 30pp; Spanish.

CC The invention relates to a chimeric antibody, or an Fv-type  
CC single-chain fragment, derived from the murine monoclonal antibody  
CC IOR C5 expressed by hybridoma ECCC 97061101. It contains the CDRs  
CC (complementarity determining regions) of IOR C5 and human constant  
CC regions, in both heavy and light chains. The IOR C5 antibody specifically  
CC binds to the IOR C2 antigen which is expressed on malignant and normal  
CC colorectal cells. The invention also encompasses a cell line which  
CC expresses the humanised IOR C5 antibody; host cells which express the  
CC Fv fragment; a pharmaceutical composition for treatment, location or  
CC in vivo identification of malignant tumours of colon and rectum,  
CC including their metastases and recurrences, containing the humanised  
CC antibody or the Fv fragment thereof, plus an excipient; and a method  
CC for the in vivo diagnosis of tumours comprising the administration of  
CC a 99m-technetium-labelled humanised IOR C5 antibody or Fv fragment, and  
CC monitoring biodistribution by radiography. The humanised IOR C5 antibody  
CC and/or its Fv fragment are used to treat malignant tumours of the colon  
CC or rectum, including their metastases or recurrences, and for in vivo  
CC diagnosis (imaging) of tumours, particularly when radiolabelled.

CC Sequences AAB73652-AAB73654 represent, respectively, CDRs 1-3 of the  
XX light chain variable region (LH) of antibody IOR C5.  
SQ Sequence 16 AA;

Query Match 72.8%; Score 59; DB 22; Length 16;  
Best Local Similarity 75.0%; Pred. No. 0.0027;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSKSKLLYKDGKTYLN 16  
:||||| |||||  
Db 1 KSSQSLSDSGKTYLN 16

RESULT 12  
AAR24704  
ID AAR24704 standard; peptide; 16 AA.  
XX AC AAR24704;  
XX DT 28-DEC-1992 (first entry)  
XX DE Sequence of an anti-human fibrin antibody light chain variable  
DE region chain designated 'A'.  
XX KW Chimeric monoclonal antibody; anti-fibrin antibody;  
KW antithrombotic agent; myocardial infarction therapy.  
XX OS Mus musculus.  
XX FN EP491351-A.  
XX PD 24-JUN-1992.  
XX PF 17-DEC-1991; 91EP-0121591.  
XX PR 18-DEC-1990; 90JP-0413829.  
XX PR 11-NOV-1991; 91JP-0294464.  
XX PA (TAKE ) TAKEDA CHEM IND LTD.  
XX PI Iwasa S, Taka H, Watanabe T, Tada H;  
XX WPI; 1992-209528/26.  
XX DR Chimeric monoclonal antibodies - contain anti-human fibrin  
PT antibody light and heavy chain variable and constant for treating  
PT thrombotic conditions e.g. myocardial infarction  
XX Claim 1; Page 49; 87pp; English.

CC The inventors claim a chimeric monoclonal antibody which contains an  
CC anti-human fibrin antibody light chain variable region contg. at  
CC least one of the polypeptide chains A, B and C  
CC (AAR24704, R24705, AAR24706) and a human antibody light chain constant  
CC region. The chimeric Abs can be used both in vivo and in vitro and, since  
CC they have very low immunogenicity as compared with mouse Abs, they can be  
CC administered to humans for diagnostic and therapeutic purposes. They are  
CC also more stable and show a longer half-life in the blood as compared  
CC with the original mouse Abs.

XX SQ Sequence 16 AA;

Query Match 70.4%; Score 57; DB 13; Length 16;  
Best Local Similarity 80.0%; Pred. No. 0.0059;  
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSKSKLLYKDGKTYLN 16  
:||||| |||||  
Db 2 SSQSLSDSGKTYLN 16

RESULT 13

AAR30448  
ID AAR30448 standard; peptide; 16 AA.  
XX  
AC AAR30448;  
XX  
DT 06-MAY-1993 (first entry)  
XX  
DE C242:11 MAb kappa chain CDR1.  
XX  
KW Kappa; chain; heavy; complementarity determining region; CDR; MAb;  
KW monoclonal antibody; C242:II; murine; IgG; hybridoma; cell line;  
KW spleen; human; colonic; adenocarcinoma; myeloma; Sp2/0; antigen;  
KW endocytosis.  
XX  
OS Synthetic.  
XX  
FN EP521842-A.  
XX  
XX 07-JAN-1993.  
XX  
PF 03-JUL-1992; 92EP-0850166.  
XX  
XX 03-JUL-1991; 91SE-0002074.  
XX  
PA (KABI ) KABI PHARMACIA AB.  
XX  
PI Holmgren J, Lind P, Lindholm L;  
XX  
XX WPI; 1993-002345/01.  
XX  
XX Monoclonal antibody reacting with CA-242 antigen - obt'd. by  
PT culturing hybridoma cell line C242:11 or mutants, useful for  
PT diagnosis and therapy of pancreatic or colorectal cancers  
XX  
XX Claim 5; Page 11; 15pp; English.  
XX  
CC The sequences given in AAR30448-50 represent the kappa chain and  
CC AAR30451-53 the heavy chain complementarity determining regions (CDR)  
CC of a monoclonal antibody (MAB), C242:II. C242:II is a monoclonal  
CC murine Ab of IgG class produced when culturing in an appropriate  
CC medium a hybridoma cell line obtained by fusing spleen cells from a  
CC mouse, which has been immunised with a human colonic adenocarcinoma  
CC cell line, with the murine myeloma cell line Sp2/0. C242:II when  
CC bound to a cell surface antigen is capable of being endocytosed or  
CC internalised into cells.  
XX  
SQ Sequence 16 AA;  
  
Query Match 67.9%; Score 55; DB 14; Length 16;  
Best Local Similarity 73.3%; Pred. No. 0.013;  
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 RSSKSLLYDKGKTYL 15  
DB 1 RSSKSLLYSHNGNTYL 15  
  
RESULT 14  
AAU07734  
ID AAU07734 standard; peptide; 16 AA.  
XX  
AC AAU07734;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Human light chain complementarity determining region, CDR1.  
XX  
KW Complementarity determining region; CDR1; Hu266; nootropic;  
KW neuroprotective; Abeta peptide; Alzheimer's disease; Down's syndrome;  
KW cerebral amyloid angiopathy; antibody; gene therapy.  
XX  
XX Homo sapiens.  
OS Synthetic.

XX  
PN WO200162801-A2.  
XX  
PD 30-AUG-2001.  
XX  
XX 26-FEB-2001; 2001WO-US06191.  
XX  
XX 24-FEB-2000; 2000US-0184601.  
PR 08-DEC-2000; 2000US-0254465.  
PR 08-DEC-2000; 2000US-0254498.  
XX  
XX (UNIW ) UNIV WASHINGTON.  
PA (ELIL ) LILLY & CO ELI.  
XX  
XX Holtzman DM, Demattos R, Bales KR, Paul SM, Tsurushita N;  
PI Vasquez M;  
PI  
XX WPI; 2001-550087/61.  
XX  
PT New humanised antibody for the treatment of Alzheimer's comprises the  
PT inhibition and reduction of the formation of amyloid plaques -  
XX  
PS Claim 14; Page 45; 63pp; English.  
XX  
CC The invention relates a humanised antibody that specifically binds  
CC an epitope contained within positions 13-28 of amyloid beta peptide,  
CC Abeta. The antibody is useful to inhibit and reduce the formation of  
CC amyloid plaques or the effects of toxic soluble Abeta species in humans  
CC the fragments are used for the manufacture of a medicament. This includes  
CC the prolonged expression of recombinant sequences of them in human  
CC tissues for the treatment of clinical/pre-clinical Alzheimer's disease,  
CC Down's syndrome or pre clinical cerebral amyloid angiopathy.  
CC Specifically, the antibody is used to sequester Abeta into plasma, brain  
CC or cerebrospinal fluid to prevent/reverse accumulation of the Abeta  
CC peptide within the brain thereby improving cognition. The present  
CC sequence is human light chain CDR1 (complementarity determining region 1)  
CC used to humanise the mouse monoclonal antibody 266 to produce Hu266.  
XX  
SQ Sequence 16 AA;  
  
Query Match 67.9%; Score 55; DB 22; Length 16;  
Best Local Similarity 62.5%; Pred. No. 0.013;  
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RSSKSLLYDKGKTYL 16  
DB 1 RSSQSLIYSDGNAYLH 16  
  
RESULT 15  
AAO14984  
ID AAO14984 standard; peptide; 16 AA.  
XX  
AC AAO14984;  
XX  
DT 05-SEP-2002 (first entry)  
XX  
DE Mouse mAb 1D9 antibody light chain complementarity determining region 1.  
XX  
KW Mouse; graft rejection; CC chemokine receptor 2 antagonist;  
KW CCR2 antagonist; anti-CCR2 antibody; kidney transplant; liver transplant;  
KW lung transplant; heart-lung transplant; pancreas transplant;  
KW bowel transplant; heart transplant; graft versus host disease;  
KW chronic graft rejection; antibody light chain; mAb 1D9; CDR1;  
KW complementarity determining region 1.  
XX  
OS Mus musculus.  
XX  
XX US2002042370-A1.  
PN  
XX 11-APR-2002.  
PD  
XX 13-APR-2001; 2001US-0835087.  
PF

XX	
PR	14-APR-2000; 2000US-0549448.
XX	(MILL-) MILLENNIUM PHARM INC.
PA	Hancock WW;
XX	
PI	WPI; 2002-351265/38.
DR	Inhibiting graft rejection, graft versus host disease or chronic
XX	rejection of a transplanted graft, involves administering a CCR2
PT	antagonist -
PT	
XX	Claim 25; Page 12; 16pp; English.
PS	The invention comprises a method of inhibiting graft rejection, graft
XX	versus host disease or chronic rejection of a transplanted graft. The
CC	method involves administering an antagonist of CC chemokine receptor 2
CC	(CCR2) and optionally an immunosuppressive agent. The CCR2 antagonist may
CC	be an anti-CCR2 antibody (i.e. containing light and heavy chain
CC	complementarity determining regions from various non-human origins). CCR2
CC	is known to be involved in the rejection of transplanted grafts. The
CC	method of the invention is useful for inhibiting graft rejection -
CC	particularly allografts such as kidney, liver, lung, heart-lung,
CC	pancreas, bowel and heart. The method of the invention is also useful for
CC	inhibiting graft versus host disease and for inhibiting chronic rejection
CC	of a transplanted graft. The present amino acid sequence represents a
CC	mouse mAb 1D9 antibody light chain complementarity determining region 1
CC	(CDR1).
XX	
SQ	Sequence 16 AA;
	Query Match 67.9%; Score 55; DB 23; Length 16;
	Best Local Similarity 68.8%; Pred. No. 0.013;
	Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy	1 RSSKSLLYDGKTYLN 16
	:   :
Db	1 KSSQSLLDSGDKTFLN 16

Search completed: April 22, 2003, 13:12:32  
Job time : 74 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003. Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:13:30 ; Search time 41 Seconds  
(without alignments)  
31.270 Million cell updates/sec

Title: US-09-674-716B-3

Perfect score: 81

Sequence: 1 RSSKSLLYKDGKTYLN 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 48364

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/prodata/1/pubpaa/US08\_NEW PUB.pbp.\*
- 2: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW PUB.pbp.\*
- 3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW PUB.pbp.\*
- 4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pbp.\*
- 5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW PUB.pbp.\*
- 6: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pbp.\*
- 7: /cgn2\_6/prodata/1/pubpaa/PCTUS\_PUBCOMB.pbp.\*
- 8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pbp.\*
- 9: /cgn2\_6/prodata/1/pubpaa/US09\_NEW PUB.pbp.\*
- 10: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pbp.\*
- 11: /cgn2\_6/prodata/1/pubpaa/US10\_NEW PUB.pbp.\*
- 12: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pbp.\*
- 13: /cgn2\_6/prodata/1/pubpaa/US60\_NEW PUB.pbp.\*
- 14: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pbp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	67.9	16	10 US-09-835-087-14	Sequence 14, Appl
2	49	60.5	16	9 US-09-518-737-8	Sequence 8, Appl
3	43	53.1	15	10 US-09-217-268B-29	Sequence 29, Appl
4	40	49.4	16	10 US-09-796-744-8	Sequence 8, Appl
5	39	48.1	16	9 US-10-091-236-14	Sequence 14, Appl
6	39	48.1	16	9 US-10-091-313-4	Sequence 4, Appl
7	39	48.1	16	9 US-10-091-268-4	Sequence 4, Appl
8	39	48.1	16	10 US-09-748-960-9	Sequence 9, Appl
9	32.5	40.1	15	9 US-09-879-461-16	Sequence 16, Appl
10	32	39.5	9	9 US-10-205-150-4	Sequence 4, Appl
11	30	37.0	16	9 US-09-876-904A-343	Sequence 343, App
12	29	35.8	10	9 US-09-809-638-497	Sequence 497, App
13	27	33.3	12	9 US-10-084-813-328	Sequence 928, App
14	27	33.3	15	9 US-10-084-813-692	Sequence 692, App
15	27	33.3	15	9 US-10-084-813-693	Sequence 693, App
16	27	33.3	15	9 US-10-084-813-694	Sequence 694, App
17	27	33.3	15	9 US-10-084-813-695	Sequence 695, App
18	26	32.1	9	9 US-10-084-813-622	Sequence 622, App
19	26	32.1	10	9 US-09-809-638-704	Sequence 704, App

20	26	32.1	12	10 US-09-850-351A-61	Sequence 61, Appl
21	26	32.1	14	9 US-09-736-968A-18	Sequence 18, Appl
22	26	32.1	14	10 US-09-736-969A-18	Sequence 18, Appl
23	26	32.1	14	10 US-09-736-960-18	Sequence 18, Appl
24	26	32.1	16	9 US-09-876-904A-344	Sequence 344, App
25	25	30.9	10	10 US-09-780-053-103	Sequence 103, App
26	25	30.9	10	10 US-09-780-053-670	Sequence 670, App
27	25	30.9	10	10 US-09-308-511-3	Sequence 3, Appl
28	25	30.9	11	9 US-09-969-037-1	Sequence 1, Appl
29	25	30.9	11	9 US-09-969-037-2	Sequence 2, Appl
30	25	30.9	12	9 US-10-084-813-927	Sequence 927, App
31	25	30.9	15	9 US-09-912-414-37	Sequence 37, Appl
32	25	30.9	15	9 US-10-161-499-26	Sequence 26, Appl
33	25	30.9	15	9 US-10-084-813-691	Sequence 691, App
34	24	29.6	12	9 US-09-818-991-9	Sequence 9, Appl
35	24	29.6	13	9 US-10-216-408-25	Sequence 25, Appl
36	24	29.6	15	10 US-09-861-294-23	Sequence 23, Appl
37	24	29.6	16	9 US-09-996-288-25	Sequence 25, Appl
38	24	29.6	16	9 US-09-996-288-106	Sequence 106, App
39	24	29.6	16	9 US-10-023-896-113	Sequence 113, App
40	24	29.6	16	10 US-09-796-848A-27	Sequence 27, Appl
41	23.5	29.0	11	9 US-09-968-561A-184	Sequence 184, App
42	23.5	29.0	11	10 US-09-192-854-105	Sequence 105, App
43	23	28.4	10	10 US-09-973-145-17	Sequence 17, Appl
44	23	28.4	12	9 US-10-057-505-26	Sequence 26, Appl
45	23	28.4	12	9 US-09-912-414-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1  
US-09-835-087-14  
; Sequence 14, Application US/09835087  
; Patent No. US20020042370A1  
; GENERAL INFORMATION:  
; APPLICANT: Wayne W. Hancock  
; TITLE OF INVENTION: Method of Treating Graft Rejection Using  
; FILE REFERENCE: 1855-2008-003  
; CURRENT APPLICATION NUMBER: US/09/835,087  
; CURRENT FILING DATE: 2001-09-24  
; \* PRIOR APPLICATION NUMBER: 09/549,448  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Mus Musculus  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)...(16)  
; OTHER INFORMATION: CDR1 of murine mAb 1D9 light chain variable region  
US-09-835-087-14

Query Match 67.9%; Score 55; DB 10; Length 16;  
Best Local Similarity 68.8%; Pred. No. 0.007;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||:|||||:  
DB 1 KSSQSLDSDGKTFN 16

RESULT 2  
US-09-518-737-8  
; Sequence 8, Application US/09518737  
; Publication NO. US20030008321A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUI, YASUHIISA  
; APPLICANT: NAGATA, SATOSHI  
; APPLICANT: SHIRAI, RYUICHI

; APPLICANT: SAITO, NAOAKI  
; TITLE OF INVENTION: MONOCLONAL ANTIBODY RECOGNIZING  
; FILE OF INVENTION: PHOSPHATIDYLINOSITOL-3,4-DIPHOSPHATE  
; FILE REFERENCE: 1965/49618  
; CURRENT APPLICATION NUMBER: US/09/518,737  
; CURRENT FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: JP 1999-250209  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-518-737-8

Query Match 60.5%; Score 49; DB 9; Length 16;  
Best Local Similarity 56.2%; Pred. No. 0.068; Length 16;  
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYL 16  
|||::: : |||  
DB 1 RSSQSLVHSGNTYL 16

## RESULT 3

US-09-217-2688-29  
; Sequence 29, Application US/092172688  
; Patent No. US20020065398A1  
; GENERAL INFORMATION:  
; APPLICANT: Mateo de Acosta del Rio, Christina M  
; APPLICANT: Rodriguez, Rolando P  
; APPLICANT: Frias, Ernesto M  
; TITLE OF INVENTION: Humanized and Chimeric Monoclonal Antibodies That Recognize Epide  
; TITLE OF INVENTION: Growth Factor Receptor (EGF-R); Diagnostic and Therapeutic Use  
; FILE REFERENCE: 2720.IUS  
; CURRENT APPLICATION NUMBER: US/09/217,2688  
; CURRENT FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Murine  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; OTHER INFORMATION: CDR of murine R3 antibody  
US-09-217-2688-29

Query Match 53.1%; Score 43; DB 10; Length 15;  
Best Local Similarity 46.7%; Pred. No. 0.62; Length 15;  
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYL 15  
|||::: : |||  
DB 1 RSSQSLVHSGNTYL 15

## RESULT 4

US-09-796-744-8  
; Sequence 8, Application US/09796744  
; Patent No. US20020098527A1  
; GENERAL INFORMATION:  
; APPLICANT: SHITARA, KENYA  
; APPLICANT: HANAI, NOBUO  
; APPLICANT: SHOJI, EMI  
; APPLICANT: SAKURADA, MIKKIKO  
; APPLICANT: FURUYA, AKIKO  
; APPLICANT: NAKAMURA, KAZUYASU  
; APPLICANT: NIWA, RINPEI  
; APPLICANT: SHIBATA, KENJI  
; APPLICANT: YAMASAKI, MOTOO  
; TITLE OF INVENTION: GENE RECOMBINANT ANTIBODY AND ANTIBODY FRAGMENT THEREOF

; FILE REFERENCE: 249-170  
; CURRENT APPLICATION NUMBER: US/09/796,744  
; CURRENT FILING DATE: 2002-01-04  
; PRIOR APPLICATION NUMBER: JP 2000-59508  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: JP 2000-401563  
; PRIOR FILING DATE: 2000-12-28  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-796-744-8

Query Match 49.4%; Score 40; DB 10; Length 16;  
Best Local Similarity 46.7%; Pred. No. 2.1; Length 16;  
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYL 15  
|||::: : |||  
DB 1 RSSRNIVHSGNTYL 15

## RESULT 5

US-10-091-236-14  
; Sequence 14, Application US/10091236  
; Patent No. US20020168360A1  
; GENERAL INFORMATION:  
; APPLICANT: DINGIVAN, CHRISTINE A.  
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE  
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING INTEGRIN ALPHA-V-BETA-3 ANTAGONISTS  
; TITLE OF INVENTION: COMBINATION WITH OTHER PROPHYLACTIC OR THERAPEUTIC AGENTS  
; FILE REFERENCE: 10271-053-999  
; CURRENT APPLICATION NUMBER: US/10/091,236  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/273,098  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: US 60/316,321  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-10-091-236-14

Query Match 48.1%; Score 39; DB 9; Length 16;  
Best Local Similarity 61.5%; Pred. No. 3; Length 16;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKT 13  
|||::: : |||  
DB 1 RSSQSLVHSGNT 13

## RESULT 6

US-10-091-313-4  
; Sequence 4, Application US/10091313  
; Publication No. US2003004406A1  
; GENERAL INFORMATION:  
; APPLICANT: DINGIVAN, CHRISTINE  
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE  
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING CD2 ANTAGONISTS IN COMBINATION WITH O  
; TITLE OF INVENTION: PROPHYLACTIC OR THERAPEUTIC AGENTS  
; FILE REFERENCE: 10271-063  
; CURRENT APPLICATION NUMBER: US/10/091,313  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/273,098  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: US 60/346,918  
; PRIOR FILING DATE: 2001-10-19

```
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent version 3.1
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-091-313-4

Query Match      48.1%; Score 39; DB 9; Length 16;
Best Local Similarity 61.5%; Pred. No. 3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKT 13
Db 1 RSSQSLHSSGNT 13

RESULT 7
US-10-091-268-4
; Sequence 4, Application US/10091268
; Publication No. US20030068320A1
; GENERAL INFORMATION:
; APPLICANT: DINGIVAN, CHRISTINE A
; TITLE OF INVENTION: METHODS OF ADMINISTERING/DOSING CD2 ANTAGONISTS FOR THE PREVENTIO
; FILE REFERENCE: 10271-054-999
; CURRENT APPLICATION NUMBER: US/10/091,268
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/273,098
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/346,918
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent version 3.1
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-091-268-4

Query Match      48.1%; Score 39; DB 9; Length 16;
Best Local Similarity 61.5%; Pred. No. 3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKT 13
Db 1 RSSQSLHSSGNT 13

RESULT 8
US-09-748-960-9
; Sequence 9, Application US/09748960
; Patent No. US20010046496A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Judith A.
; APPLICANT: Brettman, Lee R.
; APPLICANT: Allison, David Edward
; TITLE OF INVENTION: Method of Administering an Antibody
; FILE REFERENCE: 1855.2007-001
; CURRENT APPLICATION NUMBER: US/09/748,960
; CURRENT FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: US 09/550,082
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)...(16)
; OTHER INFORMATION: CDR1 of the light chain of antibodies Act-1 and
```

```
; OTHER INFORMATION: LDP-02
US-09-748-960-9

Query Match      48.1%; Score 39; DB 10; Length 16;
Best Local Similarity 56.2%; Pred. No. 3;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
Db 1 RSSQSLAKSYGNTYLS 16

RESULT 9
US-09-879-461-16
; Sequence 16, Application US/09879461
; Publication No. US2002019375A1
; GENERAL INFORMATION:
; APPLICANT: Holmes, Stephen D.
; Gross, Mitchell S.
; Sylvester, Daniel R.
; TITLE OF INVENTION: Recombinant IL4 Antibodies Useful in
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: Corporate Intellectual Property, UW2220 - 709
; Swedeland Rd.
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/879,461
; FILING DATE: 12-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,929
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/136,783
; FILING DATE: 14-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sutton, Jeffrey A.
; REGISTRATION NUMBER: 34,028
; REFERENCE/DOCKET NUMBER: P50186-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 270-5024
; TELEFAX: (215) 270-5090
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-879-461-16

Query Match      40.1%; Score 32.5; DB 9; Length 15;
Best Local Similarity 43.8%; Pred. No. 33;
Matches 7; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

QY 1 RSSKSLLYKDGKTYLN 16
Db 1 KASQSDVY-DGDSYNN 15

RESULT 10
US-10-205-150-4
; Sequence 4, Application US/10205150
```

```

; Publication No. US20020197369A1
; GENERAL INFORMATION:
; APPLICANT: LINGNAU, KAREN ET AL.
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR IMMUNOMODULATION AND PREPARATION
; TITLE OF INVENTION: OF VACCINES COMPRISING AN ANTIGEN AND AN IMMUNOGENIC OLIGODEOXYN
; FILE REFERENCE: AND A POLYCACTONIC POLYMER AS ADJUVANTS
; CURRENT APPLICATION NUMBER: SONN:018US
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/205,150
; PRIOR FILING DATE: 2002-07-25
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-205-150-4

Query Match      39.5%; Score 32; DB 9; Length 9;
Best Local Similarity 62.5%; Pred. No. 2.7e+05;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      8 YKDGKTYL 15
Db      2 YKDGNEYI 9

RESULT 11
US-09-876-904A-343
; Sequence 343, Application US/09876904A
; Publication No. US20030072794A1
; GENERAL INFORMATION:
; APPLICANT: BOULIKAS, TENI
; TITLE OF INVENTION: ENCAPSULATION OF PLASMID DNA (LIPOGENES TM) AND THERAPEUTIC
; TITLE OF INVENTION: AGENTS WITH NUCLEAR LOCALIZATION SIGNAL/FUSOGENIC PEPTIDE
; FILE REFERENCE: TB-2002.00
; CURRENT APPLICATION NUMBER: US/09/876,904A
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,925
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 343
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; OTHER INFORMATION: Ser/Thr protein kinase
US-09-876-904A-343

Query Match      37.0%; Score 30; DB 9; Length 16;
Best Local Similarity 55.6%; Pred. No. 93;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      8 YKDGKTYL 16
Db      6 YRDKHKLH 14

RESULT 12
US-09-809-638-497
; Sequence 497, Application US/09809638
; Publication No. US20030059895A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Steve Chappell Michell
; APPLICANT: Daniel E.H. Afar

```

```

; APPLICANT: Arthur B. Raitano
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 125P5C8: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.35US01
; CURRENT APPLICATION NUMBER: US/09/809,638
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 746
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 497
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-638-497

Query Match      35.8%; Score 29; DB 9; Length 10;
Best Local Similarity 55.6%; Pred. No. 81;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      8 YKDGKTYL 16
Db      2 YKEGHNYEN 10

RESULT 13
US-10-084-813-928
; Sequence 928, Application US/10084813
; Publication No. US20030068615A1
; GENERAL INFORMATION:
; APPLICANT: SAXINGER, CARL
; TITLE OF INVENTION: POLYPEPTIDES THAT BIND HIV GP120 AND RELATED NUCLEIC
; FILE REFERENCE: 215875
; CURRENT APPLICATION NUMBER: US/10/084,813
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/US00/23505
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,270
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 1242
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 928
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: binding peptide
US-10-084-813-928

Query Match      33.3%; Score 27; DB 9; Length 12;
Best Local Similarity 41.7%; Pred. No. 2.1e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      4 KSLLYKDGKTYL 15
Db      1 KNLKIEDSDTYI 12

RESULT 14
US-10-084-813-692
; Sequence 692, Application US/10084813
; Publication No. US20030068615A1
; GENERAL INFORMATION:
; APPLICANT: SAXINGER, CARL
; TITLE OF INVENTION: POLYPEPTIDES THAT BIND HIV GP120 AND RELATED NUCLEIC
; FILE REFERENCE: 215875
; CURRENT APPLICATION NUMBER: US/10/084,813
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/US00/23505
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,270
; PRIOR FILING DATE: 1999-08-27

```

```
; NUMBER OF SEQ ID NOS: 1242
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 692
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: binding peptide
US-10-084-813-692
```

```
Query Match      33.3%; Score 27; DB 9; Length 15;
Best Local Similarity 41.7%; Pred. No. 2.7e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 4 KSLLYKDGKTYL 15
   |:|:|:|:|:
Db 4 KNLKIEDSDTYI 15
```

```
RESULT 15
US-10-084-813-693
; Sequence 693, Application US/10084813
; Publication No. US20030068615A1
; GENERAL INFORMATION:
; APPLICANT: SAKINGER, CARL
; TITLE OF INVENTION: POLYPEPTIDES THAT BIND HIV GP120 AND RELATED NUCLEIC
; FILE REFERENCE: 215875
; CURRENT APPLICATION NUMBER: US/10/084,813
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/US00/23505
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,270
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 1242
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 693
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: binding peptide
US-10-084-813-693
```

```
Query Match      33.3%; Score 27; DB 9; Length 15;
Best Local Similarity 41.7%; Pred. No. 2.7e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 4 KSLLYKDGKTYL 15
   |:|:|:|:|:
Db 3 KNLKIEDSDTYI 14
```

```
Search completed: April 22, 2003, 13:21:14
Job time : 41 secs
```



GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:11:19 ; Search time 14 Seconds  
(without alignments)  
33.626 Million cell updates/sec

Title: US-09-674-716B-3  
Perfect score: 81  
Sequence: 1 RSKSKLLYDKGKTYLN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 116189

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77	95.1	16	2	US-08-672-345C-22
2	77	95.1	16	2	US-08-672-345C-25
3	77	95.1	16	4	US-09-214-095D-22
4	77	95.1	16	4	US-09-214-095D-25
5	75	92.6	16	2	US-08-672-345C-19
6	75	92.6	16	2	US-08-672-345C-28
7	75	92.6	16	4	US-09-214-095D-19
8	75	92.6	16	4	US-09-214-095D-28
9	69	85.2	16	2	US-08-672-345C-43
10	69	85.2	16	2	US-08-672-345C-79
11	69	85.2	16	4	US-09-214-095D-43
12	69	85.2	16	4	US-09-214-095D-79
13	55	67.9	16	1	US-08-438-123-1
14	52	64.2	16	4	US-09-406-532-18
15	49	60.5	16	1	US-08-244-626-4
16	46	56.8	16	1	US-08-053-171-27
17	46	56.8	16	1	US-07-977-696C-30
18	46	56.8	16	1	US-08-129-930B-30
19	46	56.8	16	4	US-08-134-346A-17
20	46	56.8	16	4	US-08-976-288A-30
21	45	55.6	16	1	US-07-977-696C-68
22	45	55.6	16	1	US-08-129-930B-68
23	45	55.6	16	4	US-08-976-288A-68
24	44	54.3	16	2	US-08-560-558E-29
25	32.5	40.1	15	2	US-08-483-636-16
26	32.5	40.1	15	2	US-08-483-632-16
27	32	39.5	9	5	PCT-US95-04975-6

28	30	37.0	9	3	US-08-159-339A-133	Sequence 133, Appl
29	30	37.0	14	1	US-08-360-125-23	Sequence 23, Appl
30	30	37.0	14	2	US-08-450-578-23	Sequence 23, Appl
31	30	37.0	14	2	US-09-017-628-23	Sequence 23, Appl
32	30	37.0	14	2	US-09-014-880-23	Sequence 23, Appl
33	30	37.0	14	4	US-08-450-363-23	Sequence 23, Appl
34	30	37.0	16	2	US-08-647-960-18	Sequence 18, Appl
35	29	35.8	12	6	5443956-15	Patent No. 5443956
36	28	34.6	15	1	US-07-720-189-11	Sequence 11, Appl
37	27	33.3	9	3	US-08-159-339A-219	Sequence 219, Appl
38	27	33.3	14	1	US-08-337-615A-16	Sequence 16, Appl
39	26	32.1	9	1	US-08-080-073-15	Sequence 15, Appl
40	26	32.1	11	2	US-08-765-536-3	Sequence 3, Appl
41	26	32.1	11	5	PCT-US95-08401-3	Sequence 3, Appl
42	26	32.1	12	1	US-07-620-669-16	Sequence 16, Appl
43	26	32.1	12	1	US-07-803-624-16	Sequence 16, Appl
44	26	32.1	12	1	US-07-998-361-16	Sequence 16, Appl
45	26	32.1	12	4	US-08-960-780-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1  
US-08-672-345C-22  
; Sequence 22, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672.345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-22

Query Match 95.1%; Score 77; DB 2; Length 16;  
Best Local Similarity 93.8%; Pred. No. 1.3e-06;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSKSKLLYDKGKTYLN 16

DB 1 RSKSKLLYDKGKTYLN 16

RESULT 2

US-08-672-345C-25  
; Sequence 25, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-25

Query Match 95.1%; Score 77; DB 2; Length 16;  
Best Local Similarity 93.8%; Pred. No. 1.3e-06;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYDKGKTYLN 16  
| | | | | : | | | | |  
Db 1 RSSKSLLYEDGKTYLN 16

RESULT 3  
US-09-214-095D-22  
; Sequence 22, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-22

Query Match 95.1%; Score 77; DB 4; Length 16;  
Best Local Similarity 93.8%; Pred. No. 1.3e-06;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYDKGKTYLN 16  
| | | | | : | | | | |  
Db 1 RSSKSLLYEDGKTYLN 16

RESULT 4  
US-09-214-095D-25  
; Sequence 25, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 25  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-25

Query Match 95.1%; Score 77; DB 4; Length 16;  
Best Local Similarity 93.8%; Pred. No. 1.3e-06;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYDKGKTYLN 16  
| | | | | : | | | | |  
Db 1 RSSKSLLYEDGKTYLN 16

RESULT 5  
US-08-672-345C-19  
; Sequence 19, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-19

Query Match 92.6%; Score 75; DB 2; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.8e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;



Qy 1 RSSKSLYKDGKTYLN 16  
|||:|||||  
pb 1 RSSRSLYRDGKTYLN 16

Best Local Similarity 87.5%; Pred. No. 2.8e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Query Match 92.6%; Score 75; DB 2; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.8e-06;  
Matches 14: Conservative 2; Mismatches 0; Indels

Qy 1 RSSKSLLYKDGKTYLN 16  
|||:|||||  
pb 1 RSSRSLLYRDGKTYLN 16

```

RESULT 7
US-09-214-095D-19
; Sequence 19, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-19

```

Query Match 92.6%; Score 757; DB 4; Length 16;

QY 1 RSSKSLYKDGKTYLN 16  
|||:|||||  
Db 1 RSSRSLYRDGKTYLN 16

```

RESULT 8
US-09-214-095D-28
; Sequence 28, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214, 095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 28
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-28

```

Query Match 92.6%; Score 75; DB 4; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.8e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels

Qy 1 RSSKSLLYKDGKTYLN 16  
|||:|||||  
pb<sup>e</sup> 1 RSSRSLLYRDGKTYLN 16

```

1  RESULT 9
2  US-08-672-345C-43
3  ; Sequence 43, Application US/08672345C
4  ; Patent No. 5948658
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Landry Donald, W.
7  ; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
8  ; NUMBER OF SEQUENCES: 108
9  ; CORRESPONDENCE ADDRESS:
10 ; ADDRESSEE: Cooper and Dunham LLP
11 ; STREET: 1185 Avenue of the Americas
12 ; CITY: New York
13 ; STATE: New York
14 ; COUNTRY: USA
15 ; ZIP: 10036
16 ; COMPUTER READABLE FORM:
17 ; MEDIUM TYPE: Floppy disk
18 ; COMPUTER: IBM PC compatible
19 ; OPERATING SYSTEM: PC-DOS/MS-DOS
20 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
21 ; CURRENT APPLICATION DATA:
22 ; APPLICATION NUMBER: US/08/672,345C
23 ; FILING DATE: 24-JUN-1996
24 ; CLASSIFICATION: 435
25 ; ATTORNEY/AGENT INFORMATION:
26 ; NAME: White, John P.
27 ; REGISTRATION NUMBER: 28,678
28 ; REFERENCE/DOCKET NUMBER: 0575/51400
29 ; TELECOMMUNICATION INFORMATION:
30 ; TELEPHONE: 212-278-0400
31 ; TELEFAX: 212-391-0525
32 ; INFORMATION FOR SEQ ID NO: 43:
33 ; SEQUENCE CHARACTERISTICS:
34 ; LENGTH: 16 amino acids
35 ; TYPE: amino acid
36 ; STRANDEDNESS: single
37 ; TOPOLOGY: linear
38 ; MOLECULE TYPE: peptide

```

## US-08-672-345C-43

Query Match 85.2%; Score 69; DB 2; Length 16;  
Best Local Similarity 81.2%; Pred. No. 2.7e-05;  
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||  
Db 1 KSSQSLLYSDGKTYLN 16

## RESULT 10

US-08-672-345C-79  
; Sequence 79, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 79:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-79

Query Match 85.2%; Score 69; DB 2; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.7e-05;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||  
Db 1 RSSKSLLYSDGKTYLN 16

## RESULT 11

US-09-214-095D-43  
; Sequence 43, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 43

; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-43

Query Match 85.2%; Score 69; DB 4; Length 16;  
Best Local Similarity 81.2%; Pred. No. 2.7e-05;  
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||  
Db 1 KSSQSLLYSDGKTYLN 16

## RESULT 12

US-09-214-095D-79  
; Sequence 79, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 79  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
; FEATURE:  
; NAME/KEY: CHAIN  
; LOCATION: (4)-(9)  
; OTHER INFORMATION: X at positions 4 and 9 represents any amino acid  
US-09-214-095D-79

Query Match 85.2%; Score 69; DB 4; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.7e-05;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
:|||||  
Db 1 RSSKSLLYSDGKTYLN 16

## RESULT 13

US-08-438-123-1  
; Sequence 1, Application US/08438123  
; Patent No. 5552293  
; GENERAL INFORMATION:  
; APPLICANT: Lindholm et al  
; TITLE OF INVENTION: TUMOR ANTIGEN SPECIFIC ANTIBODY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lowe, Price, LeBlanc & Becker  
; STREET: Suite 300, 99 Canal Center Plaza  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22314  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: DOS Text File  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/438,123  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906,350  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:

```
; NAME: J.G. Mullins
; REGISTRATION NUMBER: 33073
; REFERENCE/DOCKET NUMBER: 149-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 684 1111
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Polypeptide
US-08-438-123-1

Query Match 67.9%; Score 55; DB 1; Length 16;
Best Local Similarity 73.3%; Pred. No. 0.0056;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYL 15
   |||||.:.:.|||
Db 1 RSSKSLHSHNGNTYL 15

RESULT 14
US-09-406-532-18
; Sequence 18, Application US/09406532A
; Patent No. 6365154
; GENERAL INFORMATION:
; APPLICANT: Connie L. Erickson-Miller
; APPLICANT: Stephen D. Holmes
; APPLICANT: James D. Winkler
; TITLE OF INVENTION: TIE2 Agonist Antibodies
; FILE REFERENCE: P50843
; CURRENT APPLICATION NUMBER: US/09/406,532A
; PRIORITY FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 60/102,098
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)...(16)
; OTHER INFORMATION: light chain CDR 1
US-09-406-532-18

Query Match 64.2%; Score 52; DB 4; Length 16;
Best Local Similarity 66.7%; Pred. No. 0.018;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYL 15
   |||||.:.:.|||
Db 1 RSSKSLHSHNGNTYL 15

RESULT 15
US-08-244-626-4
; Sequence 4, Application US/08244626
; Patent No. 5502167
; GENERAL INFORMATION:
; APPLICANT: Waldmann, Herman
; APPLICANT: Walsh, Louise
; APPLICANT: Crowe, James Scott
; APPLICANT: Lewis, Alan Peter
; TITLE OF INVENTION: CDR GRAFTED HUMANISED CHIMERIC T-CELL
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, p.c.
; STREET: 555 Thirteenth Street, N. W.
```

```
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,626
; FILING DATE: July 15, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02251
; FILING DATE: December 4, 1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ernst, Barbara G.
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1808-153A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 783-6040
; TELEFAX: (202) 783-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-244-626-4

Query Match 60.5%; Score 49; DB 1; Length 16;
Best Local Similarity 56.2%; Pred. No. 0.055;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
   |||||.:.:.|||
Db 1 RSSKSLVSHNGNTYLH 16

Search completed: April 22, 2003, 13:14:20
Job time : 15 secs
```



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:20:30 ; Search time 43 seconds  
(without alignments)  
15.650 Million cell updates/sec

Title: US-09-674-716B-5

Perfect score: 31

Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 456

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	45.2	3	I50412	gene p20K protein
2	13	41.9	6	A11490	pyruvate kinase (E
3	12	38.7	5	I40469	dnazX-like protein
4	12	38.7	6	I51434	H4 histone - Afric
5	12	38.7	6	PT0518	T-cell receptor be
6	11	35.5	5	B31836	20K protein - Rick
7	11	35.5	5	H44817	34.5K structural p
8	11	35.5	5	F44817	34.5K structural p
9	11	35.5	5	A58728	serawettin W2 - S
10	11	35.5	5	S42620	aggreccan - bovine
11	10	32.3	5	A37114	hypoxanthine phosph
12	10	32.3	5	PT0295	Ig heavy chain CRD
13	10	32.3	5	PT0577	T-cell receptor be
14	10	32.3	5	PT0565	T-cell receptor be
15	10	32.3	5	PT0700	surface protein te
16	10	32.3	5	S69237	hypothetical prote
17	10	32.3	6	A37765	hypothetical prote
18	10	32.3	6	S14159	parasporal crystal
19	10	32.3	6	B26206	alpha-1,4-glucan-p
20	10	32.3	6	A46474	PC epsilon RIIIB -
21	10	32.3	7	A34026	acetylcholinestera
22	10	32.3	7	B33541	hypothetical prote
23	10	32.3	7	I55382	hypothetical pepti
24	9	29.0	3	I78890	tyrosine protein k
25	9	29.0	5	S70154	URP2 protein - Xan
26	9	29.0	5	I40702	primase - Citrobac
27	9	29.0	5	B37325	pap fibmbrial regul
28	9	29.0	5	T14910	hypothetical prote
29	9	29.0	5	PT0686	T-cell receptor be

T-cell receptor be  
dnAA protein - pse  
hypothetical 6 pro  
D-SP2.5 region - m  
T-cell receptor ga  
T-cell receptor ga  
laminin B1 - weste  
orf 3 rara 5'-regi  
microcin C7 - Esch  
piliE protein - Esc  
mabinlin II chain  
seed protein ws-5  
myomodulin - Calif  
neuromodulatory pe  
neuromodulatory pe  
neuromodulatory pe

ALIGNMENTS

RESULT 1

I50412

gene p20K protein - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C;Accession: I50412

R;Mao, P.L.; Beauchemin, M.; Bedard, P.A.

J. Biol. Chem. 268, 8131-8139, 1993

A;Title: Quiescence-dependent activation of the p20K promoter in growth-arrested chicken

A;Reference number: A46643; MUID:93216790; PMID:8463325

A;Accession: I50412

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-3 <MAO>

A;Cross-references: GB:L02537; NID:g212616; PID:g212617

C;Genetics:

A;Gene: p20K

Query Match 45.2%; Score 14; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MST 4

Db 1 MST 3

RESULT 2

A11490

pyruvate kinase (EC 2.7.1.40) - pig (fragment)

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 03-Mar-1995

C;Accession: A11490

R;Hjelmlquist, G.; Andersson, J.; Edlund, B.; Engstrom, L.

Biochem. Biophys. Res. Commun. 61, 559-563, 1974

A;Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase ph

A;Reference number: A11490; MUID:75127438; PMID:4375989

A;Accession: A11490

A;Molecule type: protein

A;Residues: 1-6 <HJE>

A;Experimental source: liver

C;Keywords: glycolysis; phosphotransferase

Query Match 41.9%; Score 13; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RAS 7

Db 3 RAS 5

```

RESULT 3
I40469
dnazX-like protein - Bacillus subtilis (fragment)
C:Species: Bacillus subtilis
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I40469
R:Struck, J.C.; Hartmann, R.K.; Toschka, H.Y.; Erdmann, V.A.
Mol. Gen. Genet. 215, 478-482, 1989
A:Title: Transcription and processing of Bacillus subtilis small cytoplasmic RNA.
A:Reference number: I40469; MUID:89218958; PMID:2468993
A:Accession: I40469
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: EMBL:X14796; NID:940130; PIDN:CAA32902.1; PID:94376204
C:Genetics:
A:Start codon: GTG

Query Match 38.7%; Score 12; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MSTR 6
DB 1 MSYQA 5

RESULT 4
I51434
H4 histone - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I51434
R:Woodland, H.R.; Warrington, J.R.; Ballantine, J.E.M.; Turner, P.C.
Nucleic Acids Res. 12, 4939-4958, 1984
A:Title: Are there major developmentally regulated H4 gene classes in Xenopus?
A:Reference number: I51391; MUID:84247348; PMID:6330691
A:Accession: I51434
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-6 <WOO>
A:Cross-references: GB:K02304; NID:g214227; PIDN:AAA49738.1; PID:g555517

Query Match 38.7%; Score 12; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MSTR 5
DB 1 MSGR 4

RESULT 5
PT0518
T-cell receptor beta chain V-D-J region (100-4B) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0518
R:Peasey, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0518
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 38.7%; Score 12; DB 2; Length 6;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 STRAS 7
DB 1 ASRAT 5

RESULT 6
B31836
20K protein - Rickettsia rickettsii (fragment)
C:Species: Rickettsia rickettsii
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 28-May-1999
C:Accession: B31836
R:Anderson, B.E.; Baumstark, B.R.; Bellini, W.J.
J. Bacteriol. 170, 4493-4500, 1988
A:Title: Expression of the gene encoding the 17-kilodalton antigen from Rickettsia rickettsii
A:Reference number: A91885; MUID:8908059; PMID:3139629
A:Accession: B31836
A:Molecule type: DNA
A:Residues: 1-5 <AND>
A:Cross-references: GB:J03371; NID:gl52455; PIDN:AA015030.1; PID:94262874

Query Match 35.5%; Score 11; DB 2; Length 5;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 MSTR 6
DB 1 MDTNS 5

RESULT 7
H44817
34.5K structural protein - Leuconostoc oenos phage P32 (fragment)
C:Species: Leuconostoc oenos phage P32
C>Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998
C:Accession: H44817
R:Arendt, E.K.; Lonvaud, A.; Hammes, W.P.
J. Gen. Microbiol. 137, 2135-2139, 1991
A:Title: Lysogeny in Leuconostoc oenos.
A:Reference number: A44817; MUID:92085033; PMID:1748868
A:Accession: H44817
A:Molecule type: protein
A:Residues: 1-5 <ARE>
A>Note: sequence extracted from NCBI backbone (NCBI:70332)

Query Match 35.5%; Score 11; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MST 4
DB 1 MAT 3

RESULT 8
F44817
34.5K structural protein - Leuconostoc oenos phage P54 (fragment)
C:Species: Leuconostoc oenos phage P54
C>Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998
C:Accession: F44817
R:Arendt, E.K.; Lonvaud, A.; Hammes, W.P.
J. Gen. Microbiol. 137, 2135-2139, 1991
A:Title: Lysogeny in Leuconostoc oenos.
A:Reference number: A44817; MUID:92085033; PMID:1748868
A:Accession: F44817
A:Molecule type: protein
A:Residues: 1-5 <ARE>
A>Note: sequence extracted from NCBI backbone (NCBI:70335)

Query Match 35.5%; Score 11; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MST 4
DB 1 MAT 3

```

```

Db      1 MAT 3

RESULT 9
A58728
Serrawettin W2 - Serratia marcescens
C:Species: Serratia marcescens
C>Date: 10-Feb-1998 #sequence_revision 12-Feb-1998 #text_change 12-Feb-1998
C:Accession: A58728
R:Macuyama, T.; Kaneda, K.; Nakagawa, Y.; Isa, K.; Hara-Hotta, H.; Yano, I.
J. Bacteriol. 174, 1769-1776, 1992
A:Title: A novel extracellular cyclic lipopeptide which promotes flagellum-dependent and
A:Reference number: A58728; MUID:92193260; PMID:1548227
A:Accession: A58728
A:Status: unencoded polypeptide
A:Molecule type: protein
A:Residues: 1-5 <MAT>
A:Experimental source: strain NS 25
C:Comment: A surfactant lipopeptide promoting flagellum-independent surface translocation
C:Keywords: blocked amino end; blocked carboxyl end; D-amino acid; lipoprotein; unencoded
F:1/Modified site: D-leucine (Leu) #status experimental
F:4/Modified site: D-phenylalanine (Phe) #status experimental
F:1-5/Cross-link: 3-hydroxydecanoyl amino end, ester carboxyl end (Leu-Ile) #status experimental

Query Match      35.5%; Score 11; DB 4; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  2 MST 4
    :|
    :|
Db   1 LST 3

RESULT 10
S42620
aggreccan - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 19-Mar-1997 #sequence_revision 24-Jul-1998 #text_change 24-Jul-1998
C:Accession: S42620
R:Vogel, K.G.; Sandy, J.D.; Pogany, G.; Robbins, J.R.
Matrix Biol. 14, 171-179, 1994
A:Title: Aggreccan in bovine tendon.
A:Reference number: S42620; MUID:94340214; PMID:7520336
A:Accession: S42620
A:Molecule type: protein
A:Residues: 1-7 <VOG>
A:Experimental source: flexor tendon
C:Keywords: cartilage

Query Match      35.5%; Score 11; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY  1 LMSTR 5
    :|
    :|
Db   3 IVSPR 7

RESULT 11
A37114
hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) - fluke (Schistosoma mansoni) (fragment)
C:Species: Schistosoma mansoni
C>Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 23-Jun-1993
C:Accession: A37114
R:Yuan, L.; Craig, S.P.; McKerrow, J.H.; Wang, C.C.
J. Biol. Chem. 265, 13528-13532, 1990
A:Title: The hypoxanthine-guanine phosphoribosyltransferase of Schistosoma mansoni. Further
A:Reference number: A37114; MUID:90337955; PMID:2199439
A:Accession: A37114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <YUA>

```

C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 32.3%; Score 10; DB 2; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MST 4  
 :|  
 :|  
 Db 1 MSS 3

RESULT 12  
 PT0295

IG heavy chain CRD3 region (clone 5-91) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0295  
 R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J  
 A:Reference number: PT0222; MUID:91108337; PMID:1899102  
 A:Accession: PT0295  
 A:Molecule type: DNA  
 A:Residues: 1-5 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 32.3%; Score 10; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TR 5  
 :|  
 :|  
 Db 3 TR 4

RESULT 13

PT0577  
 T-cell receptor beta chain V-D-J region (141-1BC) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0577; PT0574  
 R:Feeney, A.J.  
 J. Exp. Med. 174, 115-124, 1991  
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601; PMID:1711558  
 A:Accession: PT0577  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-5 <FEE>  
 A:Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1BC  
 A:Accession: PT0574  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-5 <FEE>  
 A:Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1Q  
 C:Keywords: T-cell receptor

Query Match 32.3%; Score 10; DB 2; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 STR 5  
 :|  
 :|  
 Db 2 SSR 4

RESULT 14

PT0565  
 T-cell receptor beta chain V-D-J region (141-1CF) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0565

R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0565  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-5 <FEE>  
A;Experimental source: day 19 fetal thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 32.3%; Score 10; DB 2; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 STR 5  
|:|  
Db 2 SSR 4

## RESULT 15

PT0700  
T-cell receptor beta chain V-D-J region (161-2A) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PT0700  
R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0700  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-5 <FEE>  
A;Experimental source: newborn thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 32.3%; Score 10; DB 2; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 STR 5  
|:|  
Db 2 SSR 4

Search completed: April 22, 2003, 13:24:30  
Job time : 43 secs



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:14:25 ; Search time 11 Seconds  
(without alignments)  
26.394 Million cell updates/sec

Title: US-09-674-716B-5  
Perfect score: 31  
Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 85

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	14	45.2	7	1	GFRP MOUSE
2	11	35.5	5	1	BIOA CITFR
3	11	35.5	7	1	UC24 MAIZE
4	10	32.3	6	1	VP19 HSV1K
5	9	29.0	7	1	UH11_RAT
6	9	29.0	7	1	WWA1_ACHFU
7	9	29.0	7	1	WWA2_ACHFU
8	9	29.0	7	1	WWA3_ACHFU
9	8	25.8	7	1	CIA_ENTPA
10	7	22.6	4	1	TUFT_HUMAN
11	7	22.6	7	1	IGAO_DACDE
12	6	19.4	4	1	DCMS_PSECH
13	6	19.4	4	1	FFKA_ATEL
14	6	19.4	5	1	BIOB CITFR
15	6	19.4	5	1	PRCT_PERAM
16	6	19.4	5	1	PSK_DAUCA
17	6	19.4	6	1	TEPI_PSEPT
18	6	19.4	6	1	UN06_CLOPA
19	6	19.4	7	1	CCFI_ENTPA
20	6	19.4	7	1	CHOX_ALCSP
21	5	16.1	4	1	DCML_PSECH
22	5	16.1	4	1	PAR3_HIRME
23	5	16.1	4	1	PAR4_HIRME
24	5	16.1	4	1	FLRP_HIRME
25	5	16.1	4	1	FLRN_ATEL
26	5	16.1	4	1	FMRF_MAGNI
27	5	16.1	4	1	FYRI_ATEL
28	5	16.1	4	1	RM01_YEAST
29	5	16.1	5	1	E103_LITRU
30	5	16.1	5	1	E104_LITRU
31	5	16.1	5	1	FARP_ARTTR
32	5	16.1	5	1	RE31_LITRU
33	5	16.1	5	1	RE32_LITRU

P19991 acheta dome  
P38639 mus musculus  
P25154 oryctolagus  
P13736 mytilus edu  
P13737 mytilus edu  
P82096 litoria rub  
P41966 moniezia ex  
P41491 locusta mig  
P41495 sarcophaga  
P82158 cydia pomon  
P58803 conus imper  
P10420 mytilus edu

34 5 16.1 5 1 SUGA\_ACHDO  
35 5 16.1 5 1 UF01\_MOUSE  
36 5 16.1 6 1 ACPH\_RABIT  
37 5 16.1 6 1 CIE1\_MYTED  
38 5 16.1 6 1 CIE2\_MYTED  
39 5 16.1 6 1 E101\_LITRU  
40 5 16.1 6 1 FARP\_MONEX  
41 5 16.1 6 1 LOK1\_LOCOMI  
42 5 16.1 6 1 TMOF\_SARBU  
43 5 16.1 7 1 ALL7\_CVDPO  
44 5 16.1 7 1 BRHP\_CONIM  
45 5 16.1 7 1 CARP\_MYTED

ALIGNMENTS

RESULT 1  
GFRP\_MOUSE STANDARD; PRT; 7 AA.  
AC P99025;1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).  
GN GCHFR OR GFRP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RP SEQUENCE.  
RC TISSUE=Liver;  
RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,  
RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,  
RA Cowthorne M.;  
RL Submitted (AUG-1998) to the SWISS-PROT data bank.  
CC -!- FUNCTION: MEDIATES TETRAHYDROBIOTRIN INHIBITION OF GTP  
CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE  
(BY SIMILARITY).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC SWISS-2DPAGE; P99025; MOUSE.  
FT INIT MET 0  
FT NON\_TER 7  
SQ SEQUENCE 7 AA; 806 MW; 71B5B057273B4700 CRC64;  
Query Match 45.2%; Score 14; DB 1; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.1e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMST 4  
Db 4 LIST 7

RESULT 2  
BIOA CITFR STANDARD; PRT; 5 AA.  
AC P13071;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase  
(EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA  
aminotransferase) (Fragment).  
DE aminotransferase (Fragment).  
GN BIOA.  
OS Citrobacter freundii.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Citrobacter.  
OX NCBI\_TaxID=546;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=89006280; PubMed=2971595;  
RX

```

RA Shuan D., Campbell A.;
RT "transcriptional regulation and gene arrangement of Escherichia coli,
RL Citrobacter freundii and Salmonella typhimurium biotin operons.";
CC Gene 67:203-211(1988).
CC -|- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-
CC oxononanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-
CC diaminononanoate.
CC -|- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -|- PATHWAY: Biotin biosynthesis.
CC -|- SUBUNIT: HOMODIMER.
CC -|- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21922; -; NOT ANNOTATED CDS.
CC InterPro; IPR000954; AminoTran_3.
CC DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
CC KW Biotin biosynthesis; Transferase; AminoTransferase;
CC Pyridoxal phosphate.
CC FT NON TER 5
CC SQ SEQUENCE 5 AA; 582 MW; 6AAAB1B1AGF00000 CRC64;

Query Match 35.5%; Score 11; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MST 4
| |
| |
Db 1 MTT 3

RESULT 3
UC24 MAIZE STANDARD; PRT; 7 AA.
ID UC24 MAIZE
AC P80630;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 447)
DE (Fragment)
OS Zea mays (Maize)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
CC NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -|- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.0. ITS MW IS: 30.0 kDa.
CC Maize-2DPAGE; P80630; COLEOPTILE.
CC MaizeDB; 123956; -.
CC FT NON TER 1
CC FT NON TER 7
CC SQ SEQUENCE 7 AA; 665 MW; 6DC1B5B33DC1B5D0 CRC64;

Query Match 35.5%; Score 11; DB 1; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.1e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 STRAS 7

```

```

Db 1 STAKS 5
| |
| |
RESULT 4
VP19 HSV1K STANDARD; PRT; 6 AA.
ID VP19 HSV1K
AC P23210;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Capsid assembly and DNA maturation protein (Virion protein UL38)
DE (Capsid protein VP19C) (Fragment).
DE NCBI_TaxID=10306;
GN UL38.
OS Herpes simplex virus (type 1 / strain KOS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
CC NCBI_TaxID=10306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101287; PubMed=1846198;
RA Flanagan W.M., Papavassiliou A.G., Rice M., Hecht L.B.,
RA Silverstein S., Wagner E.K.;
RT "Analysis of the herpes simplex virus type 1 promoter controlling the
RT expression of UL38, a true late gene involved in capsid assembly.";
RL J. Virol. 65:769-786(1991).
CC -|- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
CC EMBEDDED. BINDS DNA.
CC -|- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP19C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M57646; AAA5830.1; -;
CC DR Capsid assembly; Coat protein; DNA-binding.
CC FT NON TER 6
CC SQ SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;

Query Match 32.3%; Score 10; DB 1; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MST 4
| |
| |
Db 1 MKT 3

RESULT 5
UH11 RAT STANDARD; PRT; 7 AA.
ID UH11 RAT
AC P56576;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Unknown protein from 2D-page of heart tissue (Spot P11) (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar; TISSUE=Heart;
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA Jungblut P.R.;
RL Submitted (SEP-1998) to the SWISS-PROT data bank.
CC -|- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 8.5. ITS MW IS: 42 kDa.
CC FT UNSURE 2 OR A.

```

FT NON TER 7 7  
SQ SEQUENCE 7 AA; 775 MW; 6866DB040DC5A6B0 CRC64;  
Query Match 29.0%; Score 9; DB 1; Length 7;  
Best Local Similarity 66.7%; Pred. No. 1.1e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 STR 5  
DB 2 SAR 4  
RESULT 6  
WMA1\_ACHFU STANDARD; PRT; 7 AA.  
ID WMA1\_ACHFU  
AC P35919;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE WMA1-1.  
OS Achatina fulica (Giant African snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
OC Achatinacea; Achatinidae; Achatina.  
OX NCBI\_TaxID=6530;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Ganglion;  
RX MEDLINE=93265912; PubMed=8495720;  
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;  
RT "WMA1-1, -2 and -3: novel neuromodulatory peptides isolated from  
RT ganglia of the African giant snail, Achatina fulica.";  
RL FEBS Lett. 323:104-108(1993).  
CC -1- FUNCTION: EXHIBITS MODULATORY EFFECTS ON THE PERIPHERAL NERVOUS  
CC SYSTEM. INHIBITS ACTIVITY ON A CENTRAL NEURON.  
DR PIR; S33245; S33245.  
KW Neuropeptide; Amidation.  
FT MOD RES 7 7  
SQ SEQUENCE 7 AA; 993 MW; 7362D5B69B041310 CRC64;  
Query Match 29.0%; Score 9; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 MS 3  
DB 4 MS 5  
RESULT 7  
WMA2\_ACHFU STANDARD; PRT; 7 AA.  
ID WMA2\_ACHFU  
AC P35920;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE WMA2-2.  
OS Achatina fulica (Giant African snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
OC Achatinacea; Achatinidae; Achatina.  
OX NCBI\_TaxID=6530;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Ganglion;  
RX MEDLINE=93265912; PubMed=8495720;  
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;  
RT "WMA2-1, -2 and -3: novel neuromodulatory peptides isolated from  
RT ganglia of the African giant snail, Achatina fulica.";  
RL FEBS Lett. 323:104-108(1993).  
KW Neuropeptide; Amidation.  
FT MOD RES 7 7  
SQ SEQUENCE 7 AA; 964 MW; 7362D5B666D32310 CRC64;  
Query Match 29.0%; Score 9; DB 1; Length 7;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 MS 3  
DB 4 MS 5  
RESULT 8  
WMA3\_ACHFU STANDARD; PRT; 7 AA.  
ID WMA3\_ACHFU  
AC P35921;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE WMA3-3.  
OS Achatina fulica (Giant African snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
OC Achatinacea; Achatinidae; Achatina.  
OX NCBI\_TaxID=6530;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Ganglion;  
RX MEDLINE=93265912; PubMed=8495720;  
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;  
RT "WMA3-1, -2 and -3: novel neuromodulatory peptides isolated from  
RT ganglia of the African giant snail, Achatina fulica.";  
RL FEBS Lett. 323:104-108(1993).  
DR PIR; S33244; S33244.  
KW Neuropeptide; Amidation.  
FT MOD RES 7 7  
SQ SEQUENCE 7 AA; 965 MW; 7362D5B69B132310 CRC64;  
Query Match 29.0%; Score 9; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 MS 3  
DB 4 MS 5  
RESULT 9  
CIA\_ENTFA STANDARD; PRT; 7 AA.  
ID CIA\_ENTFA  
AC P11932;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Sex pheromone CAM373 (clumping-inducing agent) (CIA).  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=1351;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=87005252; PubMed=3093276;  
RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,  
RA White B.A., An F.Y., Clewell D.B., Suzuki A.;  
RT "Isolation and structure of the Streptococcus faecalis sex pheromone,  
RT CAM373.";  
RL FEBS Lett. 206:69-72(1986).  
CC -1- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS  
CC HARBORING PAM373.  
CC -1- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR  
CC SPECIFICITY OF PHEROMONES TO PLASMIDS.  
CC -1- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.  
DR PIR; A25269; A25269.  
KW Pheromone.  
FT MOD RES 7 7  
SQ SEQUENCE 7 AA; 734 MW; 75BDD72059C05DB0 CRC64;  
Query Match 25.8%; Score 8; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 6 AS 7
    ||
Db 6 AS 7

RESULT 10
TUFT_HUMAN
ID TUFT_HUMAN STANDARD; PRT; 4 AA.
AC P01858;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phagocytosis-stimulating peptide (Tuftsin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE.
RX MEDLINE=72187087; PubMed=4112769;
RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;
RT "The characteristics, isolation and synthesis of the phagocytosis
stimulating peptide tuftsin."
RL Biochem. Biophys. Res. Commun. 47:172-179(1972).
RN [2]
RP IMMUNOGLOBULIN CLASS.
RX MEDLINE=68091045; PubMed=4169272;
RA Fidalgo B.V., Najjar V.A.;
RT "The physiological role of the lymphoid system. VI. The stimulatory
effect of leucophilic gamma globulin (leucokinin) on the phagocytic
activity of human polymorphonuclear leucocyte."
RL Biochemistry 6:3386-3392(1967).
CC -!- MISCELLANEOUS: AN IGG (CALLED LEUCOKININ) BINDS REVERSIBLY TO THE
CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LEUCOKININASE ON THE
MEMBRANE RELEASES THE ACTIVE PEPTIDE TUFTSIN FROM THE GAMMA CHAIN.
CC TUFTSIN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC
ACTIVITY OF NEUTROPHILS.
DR PIR; A02147; A02147.
DR MIM; 191150; -.
SQ SEQUENCE 4 AA; 501 MW; 74176321C0000000 CRC64;

Query Match 22.6%; Score 7; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TR 5
    |
Db 1 TK 2

RESULT 11
IGAO_DACDE
ID IGAO_DACDE STANDARD; PRT; 7 AA.
AC P06234;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Galactose oxidase inhibitor.
OS Dactylium dendroideae (Cladobotryum dendroideae).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypomyces.
ON NCBI_TaxID=5132;
RN [1]_TaxID=5132;
RP SEQUENCE.
RA Avigad G., Markus Z.;
RT "Identification of a peptide inhibitor of galactose oxidase from
Dactylium dendroideae."
RL Fed. Proc. 31:447-447(1972).
CC -!- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE
GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY
BINDING TO ITS PROSTHETIC COPPER GROUP.
DR PIR; A0134; KEYDGD.

KW Copper; Metalloenzyme inhibitor.
SQ SEQUENCE 7 AA; 706 MW; 75BB01A456D87DB0 CRC64;

Query Match 22.6%; Score 7; DB 1; Length 7;
Best Local Similarity 25.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRA 6
    |
Db 4 NTES 7

RESULT 12
DCMS_PSECH
ID DCMS_PSECH STANDARD; PRT; 4 AA.
AC P19918;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO
dehydrogenase subunit S) (CO-DH S) (Fragment).
GN CUTS.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group.
ON NCBI_TaxID=290;
RN [1]_TaxID=290;
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
carboxydotrophic bacteria."
RL Arch. Microbiol. 152:335-341(1989).
CC -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
dioxide.
CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
acceptor.
CC -!- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
SMALL.
DR PIR; PLO146; PLO146.
KW Oxidoreductase; Iron-sulfur.
FT NON_TER 4
SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F0000000 CRC64;

Query Match 19.4%; Score 6; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MS 3
    |
Db 1 MA 2

RESULT 13
FFKA_ANTEL
ID FFKA_ANTEL STANDARD; PRT; 4 AA.
AC P58705;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antho-Kamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nyantheae; Actiniidae; Anthopleura.
ON NCBI_TaxID=6110;
RN [1]_TaxID=6110;
RP SEQUENCE.
RX MEDLINE=92028952; PubMed=1681803;
RA Notthacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
RT "Isolation of L-3-phenyllactyl-Phe-Lys-Ala-NH2 (Antho-Kamide), a
novel neuropeptide from sea anemones."
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
```

```

RN FUNCTION.
RP MEDLINE=93391436; PubMed=8397415;
RT McFarlane I.D., Hudman D., Nockacker H.-P., Grimmelikhuijzen C.J.P.;
RA "The expansion behaviour of sea anemones may be coordinated by two
RT inhibitory neuropeptides, Antho-Xamide and Antho-Kamide.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle
CC groups. May be involved in the expansion phase of feeding
CC behaviour in sea anemones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron-specific.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 L-3-PHENYLLACTYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 512 MW; 6DD339C9A0000000 CRC64;

Query Match 19.4%; Score 6; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RA 6
DB 3 KA 4

RESULT 14
BIOB CITFR
ID BIOR CITFR STANDARD; PRT; 5 AA.
AC P12997;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase) (Fragment).
GN BIOR.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Citrobacter.
OX NCBI_TaxID=546;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89006280; PubMed=2971595;
RA Shiuan D., Campbell A.;
RT "Transcriptional regulation and gene arrangement of Escherichia coli,
RT Citrobacter freundii and Salmonella typhimurium biotin operons.";
RL Gene 67:203-211(1988).
CC -!- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.
CC -!- PATHWAY: Biotin biosynthesis; last step.
CC -!- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21922; -; NOT ANNOTATED CDS.
KW Biotin biosynthesis; Iron-sulfur; Transferase.
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 532 MW; 75A5B1ED6F000000 CRC64;

Query Match 19.4%; Score 6; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MS 3
DB 1 MA 2

```

```

RESULT 15
PRCT PERAM STANDARD; PRT; 5 AA.
ID AC P01373;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Proctolin.
OS Periplaneta americana (American cockroach),
OS Limulus polyphemus (Atlantic horseshoe crab), and
OS Carinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattodea; Blattellidae; Periplaneta.
OX NCBI_TaxID=6978, 6850, 6759;
[1]
RP SEQUENCE.
RC SPECIES=P.americana;
RC MEDLINE=76074708; PubMed=576;
RX Starratt A.N., Brown B.E.;
RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter
RT in insects.";
RL Life Sci. 17:1253-1256(1975).
[2]
RP BIOLOGICAL SOURCE.
RC SPECIES=P.americana;
RX MEDLINE=81225865; PubMed=6113690;
RA O'Shea M., Adams M.E.;
RT "Pentapeptide (proctolin) associated with an identified neuron.";
RL Science 213:567-569(1981).
[3]
RP SEQUENCE.
RC SPECIES=L.polyphemus;
RX MEDLINE=90287800; PubMed=2356151;
RA Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
RA Shabanowitz J.;
RT "Identification of proctolin in the central nervous system of the
RT horseshoe crab, Limulus polyphemus.";
RL Peptides 11:205-211(1990).
[4]
RP SEQUENCE.
RC SPECIES=C.maenas;
RX MEDLINE=86232789; PubMed=2872661;
RA Stangier J., Dirksen H., Keller R.;
RT "Identification and immunocytochemical localization of proctolin in
RT pericardial organs of the shore crab, Carinus maenas.";
RL Peptides 7:67-72(1986).
CC -!- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY,
CC MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.
CC -!- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN
CC THE CRAB PERICARDIAL ORGANS.
DR PIR; A01644; HOROHA.
DR PIR; A60411; A60411.
KW Neuropeptide.
SQ SEQUENCE 5 AA; 649 MW; 71B7673B4600000 CRC64;

Query Match 19.4%; Score 6; DB 1; Length 5;
Best Local Similarity 33.3%; Pred. No. 1.1e+05;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MST 4
DB 3 LPT 5

Search completed: April 22, 2003, 13:22:12
Job time : 12 secs

```



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:20:00 ; Search time 81 Seconds  
(without alignments)  
17.807 Million cell updates/sec

Title: US-09-674-716B-5

Perfect score: 31

Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 58

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21.\*

- 1: sp archaea.\*
- 2: sp bacteria.\*
- 3: sp fungi.\*
- 4: sp human.\*
- 5: sp invertebrate.\*
- 6: sp mammal.\*
- 7: sp mhc.\*
- 8: sp organelle.\*
- 9: sp phage.\*
- 10: sp plant.\*
- 11: sp rodent.\*
- 12: sp virus.\*
- 13: sp vertebrate.\*
- 14: sp unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	35.5	6	10 P82541	P82541 spinacia ol
2	10	32.3	7	12 Q65578	Q65578 bovine herp
3	9	29.0	7	2 Q47505	Q47505 escherichia
4	9	29.0	7	2 P70804	P70804 azotobacter
5	9	29.0	7	8 Q95945	Q95945 saccharomyc
6	9	29.0	7	10 P93233	P93233 lycopersico
7	8	25.8	5	10 Q99007	Q99007 hordeum vul
8	7	22.6	6	10 P82181	P82181 spinacia ol
9	7	22.6	6	10 P82182	P82182 spinacia ol
10	7	22.6	7	2 Q47029	Q47029 enterobacte
11	7	22.6	7	2 Q34028	Q34028 sphingomona
12	6	19.4	5	2 P83073	P83073 bacillus ce
13	6	19.4	7	4 Q15897	Q15897 homo sapien
14	6	19.4	7	4 Q15903	Q15903 homo sapien
15	6	19.4	7	8 Q99182	Q99182 gnathotobia
16	6	19.4	7	11 Q63668	Q63668 rattus norv

17	6	19.4	7	12	Q66205	Q66205 transmissib
18	6	19.4	7	12	Q9YQ10	Q9YQ10 transmissib
19	6	19.4	7	12	Q67113	Q67113 influenzavi
20	6	19.4	7	13	Q42564	Q42564 fugu rubrip
21	6	19.4	7	15	Q07624	Q07624 rous sarcom
22	5	16.1	5	13	P83308	P83308 gallus gall
23	5	16.1	7	2	O07354	O07354 synechococc
24	5	16.1	7	2	O05556	O05556 actinobacil
25	5	16.1	7	2	Q47477	Q47477 escherichia
26	5	16.1	7	2	Q54248	Q54248 streptomyc
27	5	16.1	7	2	P72081	P72081 nocardia la
28	5	16.1	7	5	P83274	P83274 macrobrachi
29	5	16.1	7	6	Q28742	Q28742 oryctolagus
30	5	16.1	7	8	P92372	P92372 haynaldia v
31	5	16.1	7	8	P92403	P92403 lophopyrum
32	5	16.1	7	8	P92425	P92425 pseudoroegn
33	5	16.1	7	8	P92387	P92387 henrardia p
34	5	16.1	7	8	P92427	P92427 peridictyon
35	5	16.1	7	8	P92390	P92390 heteranthel
36	5	16.1	7	8	P92226	P92226 crithopsis
37	5	16.1	7	8	P92214	P92214 amblyopyrum
38	5	16.1	7	8	P92430	P92430 aegilops ta
39	5	16.1	7	8	P92221	P92221 bromus iner
40	5	16.1	7	8	P92442	P92442 taeniathecu
41	5	16.1	7	8	P92381	P92381 hordeum bra
42	5	16.1	7	8	P92393	P92393 hordeum vul
43	5	16.1	7	8	P92218	P92218 australopyr
44	5	16.1	7	8	P92440	P92440 thinopyrum
45	5	16.1	7	8	P92210	P92210 agropyron c

#### ALIGNMENTS

RESULT 1

ID P82541 PRELIMINARY; PRT; 6 AA.  
AC P82541; DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).  
OS Spinacia oleracea (Spinach).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
RC STRAIN=CV, ALVARO; TISSUE=LEAF;  
EX MEDLINE=20435797; PubMed=10874039;  
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
the small subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 37:28455-28465(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.  
CC -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.  
CC -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA  
FORM IS THE MINOR BASIC FORM.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.  
CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.  
DR InterPro; IPR002222; Ribosomal\_S19.  
DR Pfam; PF0203; Ribosomal\_S19; PARTIAL.  
DR PRINTS; PR00375; RIBOSOMALS19; PARTIAL.  
DR PROSITE; PS00323; RIBOSOMAL\_S19; PARTIAL.  
KW Ribosomal protein; Chloroplast; rRNA-binding.  
FT NON\_TER  
SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 35.5%; Score 11; DB 10; Length 6;  
Best Local Similarity 66.7%; Pred. No. 6.7e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TRA 6  
|  
|  
|  
DB 1 TRS 3

RESULT 2

ID Q65578 PRELIMINARY; PRT; 7 AA.  
AC Q65578;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical 0.8 kDa protein (Fragment).  
OS Bovine herpesvirus 1.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=10320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COOPER;  
RX MEDLINE=95313343; PubMed=7793062;  
RA Vleck C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,  
RA Letchworth G.J., Schwyzler M.;  
RT "Nucleotide sequence analysis of a 30-kb region of the bovine  
herpesvirus 1 genome which exhibits a colinear gene arrangement with  
the UL21 to UL4 genes of herpes simplex virus.";  
RL Virology 210:100-108(1995).  
DR EMBL; Z48053; CAA88130.1; -.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 758 MW; 6DD33455B1F1B1C0 CRC64;

Query Match 32.3%; Score 10; DB 12; Length 7;  
Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRA 6  
|  
|  
|  
DB 4 SNKA 7

RESULT 3

ID Q47505 PRELIMINARY; PRT; 7 AA.  
AC Q47505;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE MCCA protein.  
GN MCCA.  
OS Escherichia coli.  
OG Plasmid pmccC7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96099297; PubMed=8522520;  
RA Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;  
RT "Structure and organization of plasmid genes required to produce the  
translation inhibitor microcin C7";  
RL J. Bacteriol. 177:7131-7140(1995).  
DR EMBL; X57583; CAA40808.1; -.  
KW Plasmid.  
SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match 29.0%; Score 9; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 6.7e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MST 4  
|  
|  
|  
DB 1 MRT 3

RESULT 4

ID P70804 PRELIMINARY; PRT; 7 AA.  
AC P70804;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE ALGT protein (Fragment).  
GN ALGT.  
OS Azotobacter vinelandii.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Azotobacter.  
OX NCBI\_TaxID=354;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E;  
RX MEDLINE=96427318; PubMed=8830682;  
RA Rehm B.H.A., Ertesvag H., Valla S.;  
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algC) is  
part of an alg gene cluster physically organized in a manner similar  
to that in Pseudomonas aeruginosa.";  
RL J. Bacteriol. 178:5884-5889(1996).  
DR EMBL; X87973; CAA61230.1; -.  
KW NON\_TER 1  
SQ SEQUENCE 7 AA; 684 MW; 71B5A5A2D1AED0 CRC64;

Query Match 29.0%; Score 9; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6.7e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ST 4  
|  
|  
|  
DB 6 ST 7

RESULT 5

ID Q95945 PRELIMINARY; PRT; 7 AA.  
AC Q95945;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Inside intron 5 (Fragment).  
OS Saccharomyces cerevisiae (Baker's yeast).  
OG Mitochondrion.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=D273-10B;  
RX MEDLINE=81069885; PubMed=6254986;  
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Trzagaloff A., Macino G.;  
RT "Assembly of the mitochondrial membrane system: Structure and  
nucleotide sequence of the gene coding for subunit 1 of yeast  
cytochrome oxidase.";  
RL J. Biol. Chem. 255:11927-11941(1980).  
DR EMBL; V00694; CAA24066.1; -.  
KW Mitochondrion.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;

Query Match 29.0%; Score 9; DB 8; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6.7e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RA 6



Db ||  
1 RA 2.

RESULT 6  
P93233 PRELIMINARY; PRT; 7 AA.  
AC P93233; 29.0%; Score 9; DB 10; Length 7;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14)  
DE (Fragment).  
GN LE-ACSIB.

OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97351561; PubMed=9207843;  
RA Oetiker J.H., Olson D.C., Shiu O.Y., Yang S.F.;  
RT "Differential induction of seven 1-aminocyclopropane-1-carboxylate  
synthase genes by elicitor in suspension cultures of tomato  
(Lycopersicon esculentum)".  
RL Plant Mol. Biol. 34:275-286(1997).  
DR EMBL; U75692; AAC49682.1; -.  
KW Lyase.  
FT NON\_TER 1 1  
SQ SEQUENCE 7 AA; 828 MW; 71B412C7377415D0 CRC64;

Query Match 29.0%; Score 9; DB 10; Length 7;  
Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMST 4  
| : |  
Db 4 LVRT 7

RESULT 7  
Q99007 PRELIMINARY; PRT; 5 AA.  
AC Q99007; 29.0%; Score 9; DB 10; Length 7;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Alpha-amylase (EC 3.2.1.1) (Fragment).  
GN AMY1.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=HIMALAYA; TISSUE=ALEURONE LAYER;  
RX MEDLINE=91329704; PubMed=1831055;  
RA Jacobsen J.V., Close T.J.;  
RT "Control of transient expression of chimaeric genes by gibberellic  
acid and abscisic acid in protoplasts prepared from mature barley  
aleurone layers".  
RL Plant Mol. Biol. 16:713-721(1991).  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
CC -!- COPACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.  
CC -!- MISCELLANEOUS: THERE ARE AT LEAST 4 TYPES OF ALPHA-AMYLASE IN  
BARLEY.  
DR EMBL; X54643; CAA38455.1; -.  
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Seed; Germination;  
KW Calcium; Multigene family.  
FT NON\_TER 5 5

SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;  
Query Match 25.8%; Score 8; DB 10; Length 5;  
Best Local Similarity 25.0%; Pred. No. 6.7e+05;  
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MSTR 5  
| : |  
Db 1 MANK 4

RESULT 8  
P82181 PRELIMINARY; PRT; 6 AA.  
ID P82181  
AC P82181; 25.8%; Score 8; DB 10; Length 5;  
DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RX STRAIN=CV. ALVARO; TISSUE=LEAF;  
RX MEDLINE=20435798; PubMed=10874046;  
RA Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
the 50 S subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 275:28466-28482(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
DR InterPro; IPR001790; Ribosomal\_L10.  
DR Pfam; PF00466; Ribosomal\_L10; PARTIAL.  
DR PROSITE; PS01109; RIBOSOMAL\_L10; PARTIAL.  
KW Ribosomal protein; Chloroplast; rRNA-binding.  
FT NON\_TER 6 6  
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 22.6%; Score 7; DB 10; Length 6;  
Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TR 5  
| : |  
Db 5 TK 6

RESULT 9  
P82182 PRELIMINARY; PRT; 6 AA.  
ID P82182  
AC P82182; 22.6%; Score 7; DB 10; Length 6;  
DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RX STRAIN=CV. ALVARO; TISSUE=LEAF;  
RX MEDLINE=20435798; PubMed=10874046;  
RA Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in

RT the 50 S subunit of an organelle ribosome (chloroplast).";  
 RL J. Biol. Chem. 275:28466-28482(2000).  
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 DR InterPro: IPR001790; Ribosomal\_L10.  
 DR InterPro: IPR002363; Ribosomal\_L10eub.  
 DR Pfam: PF00466; Ribosomal\_L10; PARTIAL.  
 DR PROSITE: PS01109; RIBOSOMAL\_L10; PARTIAL.  
 DR Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON\_TER 6 6  
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;  
 Query Match 22.6%; Score 7; DB 10; Length 6;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 TR 5  
 Db 5 TK 6  
 RESULT 10  
 Q47029 PRELIMINARY; PRT; 7 AA.  
 ID Q47029; AC Q47029;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE Aad A1 protein (Fragment).  
 GN Aad A1.  
 OS Enterobacter cloacae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Enterobacter.  
 OC NCBI\_TaxID=550;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94079349; PubMed=8257126;  
 RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;  
 RT "Analysis of the aac(3)-VIA gene encoding a novel 3-N-acetyltransferase";  
 RL Antimicrob. Agents Chemother. 37:2074-2079(1993).  
 DR EMBL: M88012; AAA16193.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;  
 Query Match 22.6%; Score 7; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 TR 5  
 Db 2 TK 3  
 RESULT 11  
 Q34028 PRELIMINARY; PRT; 7 AA.  
 ID Q34028; AC Q34028;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Catechol-2,3-dioxygenase (Fragment).  
 GN PHNE.  
 OS Shingomonas chubuensis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Shingomonadaceae;  
 OC Shingomonas.  
 OC NCBI\_TaxID=56193;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RT "Isolation of chromosome-specific genes by reciprocal probing of arrayed cDNAs and cosmid libraries.";  
 RL Hum. Mol. Genet. 0:0-0(1995).  
 DR EMBL: L32077; AAA73887.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 814 MW; 672B1DD3372046B0 CRC64;  
 Query Match 19.4%; Score 6; DB 4; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;

RA Kim Y.-C.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBSJ databases.  
 DR EMBL: U88298; AAB6311.1; -.  
 KW Dioxigenase.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 868 MW; 71A452D1A699D460 CRC64;  
 Query Match 22.6%; Score 7; DB 2; Length 7;  
 Best Local Similarity 33.3%; Pred. No. 6.7e+05;  
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 MST 4  
 Db 5 VNT 7  
 RESULT 12  
 P83073 PRELIMINARY; PRT; 5 AA.  
 ID P83073; AC P83073;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE 88 kDa protein (Fragment).  
 OS Bacillus cereus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OC NCBI\_TaxID=1396;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=NCIMB 11796;  
 RA Browne N., Dowds B.C.A.;  
 RL Submitted (JUL-2001) to the SWISS-PROT data bank.  
 FT NON\_TER 5 5  
 SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;  
 Query Match 19.4%; Score 6; DB 2; Length 5;  
 Best Local Similarity 25.0%; Pred. No. 6.7e+05;  
 Matches 1; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LMST 4  
 Db 1 MKDT 4  
 RESULT 13  
 Q15897 PRELIMINARY; PRT; 7 AA.  
 ID Q15897; AC Q15897;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE (clone XP6A11A) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Lee C.-C., Yardani A., Wehnert M., Bailey J., Couch L., Xiong M.,  
 RA Coolbaugh M.I., Chinsault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,  
 RA Caskey C.T.H.;  
 RT "Isolation of chromosome-specific genes by reciprocal probing of arrayed cDNAs and cosmid libraries.";  
 RL Hum. Mol. Genet. 0:0-0(1995).  
 DR EMBL: L32077; AAA73887.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 814 MW; 672B1DD3372046B0 CRC64;  
 Query Match 19.4%; Score 6; DB 4; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;

Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RA 6  
:  
Db 4 KA 5

Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LM 2  
:  
Db 6 LL 7

## RESULT 14

Q15903  
ID Q15903 PRELIMINARY; PRT; 7 AA.  
AC Q15903;  
DT 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DE 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE (clone XP7E7B) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,  
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,  
RA Caskey C.T.H.;  
RT "Isolation of chromosome-specific genes by reciprocal probing of  
RT arrayed cDNAs and cosmid libraries.";  
RL Hum. Mol. Genet. 0:0-0(1995).  
DR EMBL; L32082; AAA73893.1; -.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 849 MW; 6B040339CDD33DB0 CRC64;

Query Match 19.4%; Score 6; DB 4; Length 7;  
Best Local Similarity 50.0%; Pred. No. 6.7e+05;

Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RA 6  
:  
Db 2 KA 3

## RESULT 15

O99182  
ID O99182 PRELIMINARY; PRT; 7 AA.  
AC O99182;  
DT 01-MAY-1999 (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DE 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE Cytochrome oxidase I (Fragment).  
GN COI.  
OS Gnatholebias zonatus.  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Gnatholebias.  
OX NCBI\_TaxID=135316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20072928; PubMed=10603257;  
RA Murphy W.J., Thomerson J.E., Collier G.E.;  
RT "Phylogeny of the Neotropical killifish family Rivulidae  
RT (Cyprinodontiformes, Aplocheiloidei) inferred from mitochondrial DNA  
RT sequences.";  
RL Mol. Phylogenet. Evol. 13:289-301(1999).  
DR EMBL; AF002591; AAD01074.1; -.  
KW Mitochondrion.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 899 MW; 672721F6CB572030 CRC64;

Query Match 19.4%; Score 6; DB 8; Length 7;  
Best Local Similarity 50.0%; Pred. No. 6.7e+05;



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003, Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:14:05 ; Search time 34 Seconds  
(without alignments)  
27.434 Million cell updates/sec

Title: US-09-674-716B-5  
Perfect score: 31  
Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 64668

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 101002:\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	7	AAW39819	Light chain CDR2 o
2	31	100.0	7	AAW39822	Light chain CDR2 o
3	31	100.0	7	AAW39825	Light chain CDR2 o
4	31	100.0	7	AAW32255	Light chain CDR L2
5	28	90.3	7	AAW39816	Light chain CDR2 o
6	27	87.1	7	AAW39876	Light chain CDR2 o
7	22	71.0	7	AAU70329	Human kappa II lig
8	19	61.3	4	RAY81430	Amino acids encode
9	19	61.3	7	AAW06840	Peptide Seq ID No:
10	19	61.3	7	AAW05015	Tumour antigen ant

11	18	58.1	6	22	AAW35205	Retroviral recomb
12	18	58.1	7	15	AAW46961	Peptide fused to i
13	18	58.1	7	16	AAW76083	Mab 55.1 light cha
14	18	58.1	7	18	AAW31179	Hepatitis C virus
15	18	58.1	7	20	AAW05016	Tumour antigen ant
16	18	58.1	7	21	AAW95223	Anti-platelet glyc
17	17	54.8	6	20	AAW95268	Anti-progesterone
18	17	54.8	7	14	AAW30157	Mab GAH variable r
19	17	54.8	7	15	AAW57964	Mab NPS2 light cha
20	17	54.8	7	15	AAW57966	Mab NPS2 light cha
21	17	54.8	7	15	AAW50317	Mab NPS2 light cha
22	17	54.8	7	15	AAW54108	Humanised anti-Hiv
23	17	54.8	7	17	AAW03579	G-protein coupled
24	17	54.8	7	17	AAW98485	Anti-IL-5 Mab ligh
25	17	54.8	7	19	AAW68514	Mab MCP603 V1 CDR
26	17	54.8	7	19	AAW42461	Mouse anti-human I
27	17	54.8	7	20	AAW08713	Human milk fat glo
28	17	54.8	7	21	AAW07952	CDR2 sequence from
29	17	54.8	7	21	AAW95234	Human monoclonal a
30	17	54.8	7	21	AAW54529	Human CD4 protein
31	17	54.8	7	21	AAW32227	G-CSF agonist anti
32	17	54.8	7	22	AAW62866	Anti-SAF-1 antibod
33	17	54.8	7	23	ABG68854	CDR2 V1 peptide.
34	17	54.8	7	23	AAU81257	Human trkC antibod
35	17	54.8	7	23	AAU70335	Human kappa IV lig
36	17	54.8	7	23	AAU70347	Mouse kappa I ligh
37	16	51.6	4	12	AAW14267	Chromogenic substr
38	16	51.6	4	20	AAW30776	Peptide used to de
39	16	51.6	5	21	AAW19563	Human matrixphage s
40	16	51.6	6	19	AAW57028	Enzyme inhibitor p
41	16	51.6	6	19	AAW56931	Enzyme inhibitor p
42	16	51.6	6	21	AAW24306	Prostate tumour as
43	16	51.6	6	22	AAW55470	Human elastase var
44	16	51.6	7	17	AAW07065	Synthetic peptide
45	16	51.6	7	19	AAW70909	CDR2 of the light

ALIGNMENTS

RESULT 1

AAW39819

ID AAW39819 standard; peptide; 7 AA.

AC AAW39819;

XX

XX

DT 16-JUN-1998 (first entry)

DE Light chain CDR2 of catalytic antibody 6A12.

XX

XX

KW Variable domain; lambda light chain; catalytic antibody; degradation;

KW cocaine; cocaine transition state analogue; TSA; benzoic acid;

KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;

KW overdose; addiction.

XX

OS Mus sp.

XX

XX

PN WO9749800-A1.

XX

PD 31-DEC-1997.

XX

XX

PF 25-JUN-1997; 97WO-US10965.

XX

XX

PR 25-JUN-1996; 96US-0672345.

XX

XX

PA (UYCO ) UNIV COLUMBIA NEW YORK.

XX

PI Landry DW;

XX

XX

DR WPI; 1998-077166/07.

XX

XX

PT New catalytic antibodies able to decompose cocaine, single-chain

PT analogues - used to treat cocaine overdose and addiction, required

PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding

PS Claim 13; Page 81; 147pp; English.

XX AAW39818-20 represent the sequences of the light chain complementarity  
 CC determining regions (CDRs) of the catalytic antibody 6A12, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody was  
 CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
 CC monoester transition state analogue. Antibody 6A12 has a per minute Kcat  
 CC of 0.072. The antibodies reduce the concentration of cocaine in a  
 CC subject, and are used particularly for the treatment of an overdose. They  
 CC are also used for treating addiction (by reducing the in vivo  
 CC concentration that can be achieved).

SQ Sequence 7 AA;

Query Match 100.0%; Score 31; DB 19; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

DB 1 LMSTRAS 7

RESULT 2

AAW39822  
 ID AAW39822 standard; peptide; 7 AA.

AC AAW39822;

XX 16-JUN-1998 (first entry)

DE Light chain CDR2 of catalytic antibody 2A10.

XX Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.

OS Mus sp.

XX WO9749800-A1.

XX 31-DEC-1997.

XX 25-JUN-1997; 97WO-US10965.

XX 25-JUN-1996; 96US-0672345.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Landry DW;

XX WPI; 1998-077166/07.

XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding

XX Claim 15; Page 82; 147pp; English.

XX AAW39821-23 represent the sequences of the light chain complementarity  
 CC determining regions (CDRs) of the catalytic antibody 2A10, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 2A10 antibody was

CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
 CC monoester transition state analogue. Antibody 2A10 has a per minute Kcat  
 CC of 0.011. The antibodies reduce the concentration of cocaine in a  
 CC subject, and are used particularly for the treatment of an overdose. They  
 CC are also used for treating addiction (by reducing the in vivo  
 CC concentration that can be achieved).

XX Sequence 7 AA;

Query Match 100.0%; Score 31; DB 19; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

DB 1 LMSTRAS 7

RESULT 3

AAW39825  
 ID AAW39825 standard; peptide; 7 AA.

XX AC AAW39825;

XX 16-JUN-1998 (first entry)

DE Light chain CDR2 of catalytic antibody 12H1.

XX Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.

OS Mus sp.

XX WO9749800-A1.

XX 31-DEC-1997.

XX 25-JUN-1997; 97WO-US10965.

XX 25-JUN-1996; 96US-0672345.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Landry DW;

XX WPI; 1998-077166/07.

XX New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding

XX Claim 17; Page 83; 147pp; English.

XX AAW39824-26 represent the sequences of the light chain complementarity  
 CC determining regions (CDRs) of the catalytic antibody 12H1, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody was  
 CC identified using TSA2, and has a per minute Kcat of 0.16. The antibodies  
 CC reduce the concentration of cocaine in a subject, and are used  
 CC particularly for the treatment of an overdose. They are also used for  
 CC treating addiction (by reducing the in vivo concentration that can be  
 CC achieved).

XX Sequence 7 AA;

Query Match 100.0%; Score 31; DB 19; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
Db 1 LMSTRAS 7

RESULT 4

AAV32255  
ID AAV32255 standard; Peptide; 7 AA.

AC AAV32255;

DT 15-FEB-2000 (first entry)

DE Light chain CDR L2 of mouse anti-CD23 Mab C11.

KW CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse;  
KW monoclonal antibody; chimeric antibody; humanised antibody;  
KW complementarity determining region; CDR; autoimmune disease;  
KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
KW urticaria; nephrotic syndrome; glomerulonephritis;  
KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
KW graft-versus-host disease; COPD; bronchitis; diabetes;  
KW B-cell malignancy; therapy.

OS Mus musculus.

PN WO9958679-A1.

XX 18-NOV-1999.

PF 07-MAY-1999; 99WO-GB01434.

PR 09-MAY-1998; 98GB-0009839.

XX (GLAX) GLAXO GROUP LTD.

PI Bonney J, Crowe SJ, Ellis JH, Rapson NT, Shearin J;

DR WPI; 2000-053101/04.

XX N-PSDB; AAZ34749.

PT Cell receptor specific antibodies useful for treating e.g. arthritis,  
PT diabetes, multiple sclerosis and psoriasis -

XX Claim 1; Page 40; 81pp; English.

XX This sequence represents complementarity determining region 2  
CC (CDR L2) of the light chain of murine anti-CD23 (FCERII) monoclonal  
CC antibody C11 (see also AAV32262). The invention provides altered  
CC antibodies, such as chimeric or humanised antibodies, which comprise  
CC sufficient of the amino acid sequences of C11 light and heavy chain  
CC CDRs (see AAV32254-59) to render them capable of binding to the CD23  
CC type II molecule expressed on haematopoietic cells. The antibodies  
CC are used to block soluble CD23 formation for treatment of arthritis,  
CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
CC (particularly chronic bronchitis) or diabetes (particularly type 1  
CC diabetes), and B-cell malignancies (claimed). They are also useful  
CC for studying interactions between CD23 and various ligands and  
CC determining the binding agents.

XX Sequence 7 AA;

Query Match 100.0%; Score 31; DB 21; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
Db 1 LMSTRAS 7

RESULT 5

AAW39816

ID AAW39816 standard; peptide; 7 AA.

XX AC AAW39816;

DT 16-JUN-1998 (first entry)

DE Light chain CDR2 of catalytic antibody 3B9.

KW Variable domain; lambda light chain; catalytic antibody; degradation;  
KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
KW overdose; addiction.

OS Mus sp.

PN WO9749800-A1.

XX 31-DEC-1997.

PF 25-JUN-1997; 97WO-US10965.

PR 25-JUN-1996; 96US-0672345.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Landry DW;

DR WPI; 1998-077166/07.

XX New catalytic antibodies able to decompose cocaine, single-chain  
XX analogues - used to treat cocaine overdose and addiction, required  
XX in far smaller doses than antibodies that antagonise cocaine by  
XX simply binding

PS Claim 11; Page 80; 147pp; English.

XX AAW39815-17 represent the sequences of the light chain complementarity  
CC determining regions (CDRs) of the catalytic antibody 3B9, which is able  
CC to degrade cocaine. A series of cocaine transition state analogues  
CC (TSAs) were prepared and used to immunise mice for production of  
CC hybridomas. Catalytic antibodies were identified by their capacity to  
CC release 3H-benzoic acid from 3H-phenyl cocaine. The 3B9 antibody was  
CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
CC monoester transition state analogue. Antibody 3B9 has a per minute Kcat  
CC of 0.11. The antibodies reduce the concentration of cocaine in a subject,  
CC and are used particularly for the treatment of an overdose. They are also  
CC used for treating addiction (by reducing the in vivo concentration that  
CC can be achieved).

XX Sequence 7 AA;

Query Match 90.3%; Score 28; DB 19; Length 7;

Best Local Similarity 85.7%; Pred. No. 7.8e+05;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7

Db 1 LMSTRSS 7

RESULT 6

AAW39876

ID AAW39876 standard; peptide; 7 AA.

XX AC AAW39876;

```

XX DT 16-JUN-1998 (first entry)
XX DE Light chain CDR2 of a catalytic antibody capable of degrading cocaine.
XX KW Variable domain; lambda light chain; catalytic antibody; degradation;
XX KW cocaine; cocaine transition state analogue; TSA; benzoic acid;
XX KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;
XX KW overdose; addiction.
XX OS Mus sp.
XX FH Key Location/Qualifiers
XX FT Misc-difference 6
XX FT /note= "not specified"
XX PN WO9749800-A1.
XX PD 31-DEC-1997.
XX PF 25-JUN-1997; 97WO-US10965.
XX PR 25-JUN-1996; 96US-0672345.
XX PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX PI Landry DW;
XX DR WPI; 1998-077166/07.
XX PT New catalytic antibodies able to decompose cocaine, single-chain
XX PT analogues - used to treat cocaine overdose and addiction, required
XX PT in far smaller doses than antibodies that antagonise cocaine by
XX PT simply binding
XX PS Claim 10; Page 5; 147pp; English.
XX CC AA039875-77 represent the sequences of the light chain complementarity
XX CC determining regions (CDRs) of a catalytic antibody which is capable of
XX CC degrading cocaine. A series of cocaine transition state analogues (TSAs)
XX CC were prepared and used to immunise mice for production of hybridomas.
XX CC Catalytic antibodies were identified by their capacity to release
XX CC 3H-benzoic acid from 3H-phenyl cocaine. These antibodies were found
XX CC to have CDRs of the present sequence. The antibodies reduce the
XX CC concentration of cocaine in a subject, and are used particularly for the
XX CC treatment of an overdose. They are also used for treating addiction (by
XX CC reducing the in vivo concentration that can be achieved).
XX SQ Sequence 7 AA;
    Query Match 87.1%; Score 27; DB 19; Length 7;
    Best Local Similarity 85.7%; Pred. No. 7.8e+05;
    Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
DB 1 LMSTRXS 7

RESULT 7
AAU070329
ID AAU070329 standard; Peptide; 7 AA.
XX AC AAU070329;
XX DT 14-FEB-2002 (first entry)
XX DE Human Kappa II light chain CDR2.
XX KW Immunoglobulin; antibody; light chain; heavy chain; CDR; FR;
XX KW complementarity determining region; framework region; IgBP;
XX KW transgenic plant; immunoglobulin binding protein array;
XX KW IgM; IgG; IgA; IgD; IgE; IgG; IgM; kappa; lambda; CHBP.

```

```

XX OS Homo sapiens.
XX PN WO200183806-A1.
XX PD 08-NOV-2001.
XX PF 02-MAY-2001; 2001WO-US14349.
XX PR 02-MAY-2000; 2000US-0563222.
XX PA (EPIC-) EPICYTE PHARM INC.
XX PI Hiatt AC, Hein WB;
XX DR WPI; 2002-055482/07.
XX PT Preparing immunoglobulin binding protein array in plant cells by
XX PT transforming the cells with different polynucleotides encoding binding
XX PT protein polypeptides specific to ligand, selecting plant cells for
XX PT preparing array
XX PS Disclosure; Page 14; 129pp; English.
XX CC The invention relates to transforming a population of cells (e.g. plant
XX CC cells), comprising using a library of two different polynucleotides
XX CC encoding different immunoglobulin binding protein (IgBP) polypeptides
XX CC that specifically bind to a ligand or form one or more disulphide bonds
XX CC with polypeptides in transfected cells, to generate an IgBP that
XX CC binds to a ligand, and transformed plant cells are selected, and
XX CC preparing an IgBP array in plant cells. At least one peptide sequence has
XX CC at least 75% sequence identity to a framework region (FR) of a native
XX CC IgM, IgG, IgA, IgD, IgE, IgY, kappa or lambda immunoglobulin molecule.
XX CC The method is useful for preparing an immunoglobulin binding protein
XX CC array, preferably heavy chain binding protein (CHBP) array in eukaryotic
XX CC cells especially plant cells (e.g. Agrobacterium tumefaciens or maize)
XX CC or other eukaryotic cells (e.g. insect cells or mammalian cells). The
XX CC CHBP is useful for discovery of e.g. screening assays of IgBPs having
XX CC desired characteristics. The present sequence is a mammalian
XX CC immunoglobulin derived peptide that may be incorporated into an IgBP of
XX CC the invention.
XX SQ Sequence 7 AA;
    Query Match 71.0%; Score 22; DB 23; Length 7;
    Best Local Similarity 71.4%; Pred. No. 7.8e+05;
    Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
DB 1 LVSNRAS 7

RESULT 8
AAU81430
ID AAU81430 standard; peptide; 4 AA.
XX AC AAU81430;
XX DT 19-JUN-2000 (first entry)
XX DE Amino acids encoded by 5' portion of pTugAS vector SacI-HindIII region.
XX KW pTugAS vector; Cellulomonas fimi; endoglucanase C; CBD;
XX KW cellulose binding domain; polysaccharide binding peptide;
XX KW fusion protein; recombinant protein; cell separation;
XX KW affinity phase separation.
XX OS Synthetic.
XX PN US6048715-A.
XX PD 11-APR-2000.

```



XX PF 24-JUL-1996; 96US-0685808.  
 XX PR 08-JUL-1988; 88US-0216794.  
 XX PR 08-APR-1992; 92US-0865095.  
 XX PR 25-OCT-1990; 90US-0603987.  
 XX PR 24-MAY-1994; 94US-0249037.  
 XX PR 24-JUL-1995; 95US-0505860.  
 XX PA (UYBR-) UNIV BRITISH COLUMBIA.  
 XX PI Haynes CA, Tomme P, Kilburn DG;  
 XX XX WPI; 2000-328038/28.  
 XX XX Two-phase partition affinity separation system useful for separating  
 PT and purifying proteins comprises a phase-forming oligosaccharide  
 PT polymer and a phase-separation agent -  
 XX XX Disclosure; Fig 3A; 46pp; English.  
 XX XX The invention relates to a novel two-phase partition system for affinity  
 CC separation which comprises a phase-forming oligosaccharide polymer to  
 CC which a polysaccharide binding peptide (PBP) binds, and a phase  
 CC separation inducing agent. The system may be used for the separation and  
 CC purification of recombinant proteins from cell cultures. The protein of  
 CC interest is expressed in the host cell as a fusion protein with a  
 CC polysaccharide binding peptide, and the cell lysate or fermentation  
 CC broth (or other composition comprising the fusion protein) is contacted  
 CC with the phase separation system. The fusion protein partitions into the  
 CC oligosaccharide polymer phase, and can then be isolated e.g., via the  
 CC use of a solution with a low ionic strength, high pH or containing a  
 CC chaotropic agent, or by the use of chemical cleavage agents such as  
 CC cyanogen bromide. If the fusion protein has been engineered such that  
 CC there is a protease recognition site between the PBP and the protein of  
 CC interest, the protein can be cleaved from the PBP, which remains bound  
 CC to the oligosaccharide. Proteins that may be purified using the method  
 CC of the invention include a wide variety of medically and industrially  
 CC important proteins e.g., interleukin 2, factor X, insulin, ligninase.  
 CC The system can also be used for cell separation and/or enrichment of a  
 CC particular cell type; for example, a fusion protein comprising a PBP and  
 CC a ligand of a particular receptor can be used to isolate cells  
 CC comprising that receptor. The system is useful for concentrating a  
 CC component in a mixture, removing contaminants and for preparing solid  
 CC state reagents for diagnostic assays. Oligosaccharide polymers can be  
 CC obtained inexpensively and the water-soluble cellulosic substrates offer  
 CC a new, cost-effective, highly flexible affinity partition system for  
 CC continuous purification. Selective binding of PBP from the  
 CC oligosaccharide polymer makes it suitable for purification of a wide  
 CC variety of compounds using a single oligosaccharide polymer phase  
 CC separation system. Hence it is unnecessary to prepare separate systems  
 CC for each compound to be separated. Sequences AAY81430-181431 represent  
 CC the amino acids sequences encoded by the 5' and 3' portions of the  
 CC SacI-HindIII region of the pTUGAS vector. DNA encoding the Cellulomonas  
 CC fimi endoglucanase C cellulose binding domain (CBD) N1 (AAY81423) was  
 CC inserted between the 5' and 3' portions of the SacI-HindIII region of the  
 CC vector.  
 XX SQ Sequence 4 AA;  
 Query Match 61.3%; Score 19; DB 21; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 MSTR 5  
 Db 1 MSTR 4  
 RESULT 9  
 AAY06840  
 ID AAY06840 standard; peptide; 7 AA.  
 XX

AC AAY06840;  
 XX 25-JUN-1999 (first entry)  
 DT Peptide Seq ID No:6 of JP11089576.  
 DE AntiHBS; monoclonal antibody; Epstein Barr virus; EBV; adr type;  
 XX human, HBS antigen, hepatitis C.  
 KW Homo sapiens.  
 OS JPI1089576-A.  
 XX 06-APR-1999.  
 PD 19-SEP-1997; 97JP-0255705.  
 XX 19-SEP-1997; 97JP-0255705.  
 PR (NISN ) NISSHINBO IND INC.  
 XX WPI; 1999-281053/24.  
 DR Anti-HBS monoclonal antibody - produced without the risk of Epstein  
 XX Barr virus contamination  
 PT Claim 4; Page 7; 12pp; Japanese.  
 PS The invention relates to an antiHBS monoclonal antibody having the  
 CC following properties: (A) CDR-3 of H chain variable region; (B) it  
 CC contains no Epstein Barr virus (EBV); (C) it binds at least one adr type  
 CC among human HBS antigens. The antiHBS monoclonal antibody is high in  
 CC antibody titer and has low risk of EBV contamination. It can be used to  
 CC prevent hepatitis C.  
 XX SQ Sequence 7 AA;  
 Query Match 61.3%; Score 19; DB 20; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 STRAS 7  
 Db 3 STRAT 7  
 RESULT 10  
 AAY05015  
 ID AAY05015 standard; peptide; 7 AA.  
 XX AAY05015;  
 AC 16-JUN-1999 (first entry)  
 XX Tumour antigen antibody light chain CDR2 clone F14.  
 DE Tumour antigen; antibody; CDR; complementarity determining region;  
 XX binding molecule identification; tumour-specific binding polypeptide;  
 KW cancer therapy; light chain.  
 OS Homo sapiens.  
 XX WO9906834-A2.  
 PN 11-FEB-1999.  
 PD 04-AUG-1998; 98WO-US16280.  
 XX 04-AUG-1997; 97US-0905825.  
 PR 04-AUG-1997; 97US-1112222.  
 XX (IXSY-) IXSYS INC.  
 XX

PI Huse WD, Watkins JD, Wu H;  
 XX WPI; 1999-153951/13.  
 XX Identifying binding molecules for ligands, particularly tumour  
 XX antigens - by selectively immobilising a population of binding  
 PT molecules to a solid support and screening for binding to two or  
 PT more ligands  
 XX Claim 15; Page 57; 80pp; English.  
 XX This sequence represents a light chain complementarity determining  
 CC region (CDR) from a tumour antigen specific antibody.  
 CC The invention relates to a method for identifying a binding molecule  
 CC having selective affinity for a ligand comprising: (a) selectively  
 CC immobilising a diverse population of binding molecules to a solid  
 CC support; (b) simultaneously contacting the diverse population immobilised  
 CC on the solid support with 2 or more ligands; and (c) determining at least  
 CC one binding molecule which selectively binds to one or more of the  
 CC ligands. The method allows for the rapid and efficient methods for the  
 CC identification of binding molecules which exhibit selective affinity for  
 CC one or more ligands of interest. They are used particularly for  
 CC identifying tumour-specific binding polypeptides which can be used as  
 CC targeting agents for cancer therapy that minimises impact on non-tumour  
 CC tissues.  
 XX SQ Sequence 7 AA;  
 Query Match 61.3%; Score 19; DB 20; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 STRAS 7  
 DB 3 STRAT 7  
 RESULT 11  
 AAB35205  
 ID AAB35205 standard; Peptide; 6 AA.  
 AC AAB35205;  
 DT 24-APR-2001 (first entry)  
 DE Retroviral recombination assay peptide fragment #3.  
 XX Retroviral recombination assay; gene therapy vector; viral vector;  
 KW gag; pol; replication; HIV.  
 XX Unidentified.  
 OS WO200104360-A2.  
 PN 18-JAN-2001.  
 PD 06-JUL-2000; 2000WO-US18597.  
 PF 09-JUL-1999; 99US-0143015.  
 PR 10-NOV-1999; 99US-0164626.  
 XX (UABR-) UAB RES FOUND.  
 PA (TRAN-) TRANZYME INC.  
 XX Kappes JC, Wu X, Wakefield J;  
 PI WPI; 2001-091927/10.  
 DR N-PSDB; AAF24358.  
 XX Retroviral recombination assays, systems and cells, useful for  
 PT evaluating the risk of producing a replication-competent retrovirus  
 PT from a retroviral-based vector -  
 XX

PS Example 5; Fig 14; 88pp; English.  
 XX The present invention describes a method for detecting a retroviral  
 CC genetic recombinant having gag and pol functions, involving a cell  
 CC suspected of having a recombinant and propagating the recombinant in the  
 CC presence of any necessary helper functions. This is useful as it enables  
 CC the detection of vectors which are able to replicate in the host cell,  
 CC and allows the production of vectors suitable for gene therapy. The  
 CC present sequence is a peptide fragment used to demonstrate the method of  
 CC the invention.  
 XX SQ Sequence 6 AA;  
 Query Match 58.1%; Score 18; DB 22; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 MSTRAS 7  
 DB 1 MGRAS 6  
 RESULT 12  
 AAR46961  
 ID AAR46961 standard; peptide; 7 AA.  
 XX AAR46961;  
 AC AAR46961;  
 DT 25-OCT-1994 (first entry)  
 DE Peptide fused to influenza A polypeptide for use in vaccine.  
 XX Peptide; influenza; vaccine; virus; antigen; expression; cleavage;  
 KW fusion protein.  
 XX Synthetic.  
 OS WO9406468-A.  
 PN 31-MAR-1994.  
 PD 17-SEP-1992; 92WO-US07312.  
 PF 17-SEP-1992; 92WO-US07312.  
 PR (SMK ) SMITHKLINE BEECHAM CORP.  
 PA Dillon SB, Jones CS, Scott MO, Shatzman A;  
 PI WPI; 1994-118165/14.  
 DR Vaccine compan. for immunising against influenza A - comprises a  
 PT portion of a haemagglutinin 2 sub unit peptide of an influenza  
 PT sub-type  
 XX Claim 11; Page 45; 58pp; English.  
 XX This sequence may be fused to the haemagglutinin 2 (HA2) fragment  
 CC (amino acids 66-222) of influenza A and the resulting fusion protein  
 CC used in a vaccine. The purpose of this particular sequence is not  
 CC given in the specification although such additional peptides to the  
 CC virus antigen may contain another influenza antigen or other antigen  
 CC or confer some advantage upon expression of the HA2 peptide.  
 CC Alternatively the fusion peptide may comprise a sequence which is  
 CC selectively cleavable by a proteolytic enzyme or upon exposure to a  
 CC selected chemical.  
 XX SQ Sequence 7 AA;  
 Query Match 58.1%; Score 18; DB 15; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRA 6  
| | | | :  
Db 2 LSTRES 7

## RESULT 13

AAAR76083  
ID AAR76083 standard; Peptide; 7 AA.

XX AC AAR76083;

XX DT 21-NOV-1995 (first entry)

XX DE MAB 55.1 light chain CDR2.

XX KW Antigen binding structure; complementarity determining region; CDR;  
KW CA55.1; colorectal cancer; tumor-associated antigen; hybridoma;  
KW monoclonal antibody; Mab; immunotherapy; therapy; diagnosis;  
KW transgenic animal; transgenic plant; antibody engineering;  
KW humanized antibody; immunotoxin.

XX OS Mus sp.

XX XX WO9515382-A.

XX PD 08-JUN-1995.

XX PF 29-NOV-1994; 94WO-GB02610.

XX PR 03-JUN-1994; 94GB-0011089.

XX PR 03-DEC-1993; 93GB-0024819.

XX PA (ZENE ) ZENECA LTD.

XX PI Blakey DC, Boot C, Copley CG, Hall SM, Paterson DS;

XX PI Rose MS, Wright AF;

XX XX WPI; 1995-215262/28.

XX DR Antigen binding structures containing CDRs recognising the CA55.1  
PT antigen - produced by hybridomas and host cells, for use in the  
PT diagnosis and therapy of cancer

XX PS Claim 2; Page 97; 121pp; English.

XX CC An antigen binding structure is based on the CDRs (given in AAR76078-  
CC 84) of the heavy and light chains of Mab 55.1 (ECACC 93081901),  
CC which recognises the colorectal tumor-associated antigen CA55.1.  
CC It is optionally humanized and in the form F(ab')<sub>2</sub>, F(ab)', Fab,  
CC Fv, scFv or V-min, and is produced in transgenic animals or plants.

XX SQ Sequence 7 AA;

Query Match 58.1%; Score 18; DB 16; Length 7;  
Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRAS 7  
| | | | :  
Db 3 STRTS 7

## RESULT 14

AAW31179

ID AAW31179 standard; peptide; 7 AA.

XX AC AAW31179;

XX DT 05-FEB-1998 (first entry)

XX DE Hepatitis C virus genotype 5a peptide, OMM25.

XX KW Hepatitis C virus; HCV; genotype determination; 1a; 1b; 2a; 2b; 3a;

KW 3b; 4; 5a; 6a; 6b; diagnosis; amplification; PCR; primer.  
XX Hepatitis C virus.

XX PN JP09234072-A.

XX PD 09-SEP-1997.

XX PF 01-FEB-1996; 96JP-0038875.

XX PR 30-DEC-1995; 95JP-0352511.

XX PR 01-FEB-1995; 95JP-0035997.

XX PA (SRLS-) SRL KK.

XX XX WPI; 1997-497313/46.

XX PT Primers used for determining hepatitis C virus genotype - provide a  
PT rapid and accurate method of hepatitis C virus genotyping  
XX Example 1; Page 18; 33pp; Japanese.

XX CC AAW31170-W31181 are peptides derived from various genotypes of  
CC hepatitis C virus. The peptides were used for the production of PCR  
CC (polymerase chain reaction) primers used for discrimination between  
CC hepatitis C virus (HCV) genotypes. Classification of the  
CC genotype of HCV can be achieved precisely and simply according to  
CC the international standardisation of classification. The primers  
CC can be used to distinguish between HCV genotypes 1a, 1b, 2a, 2b,  
CC 3a, 3b, 4, 5a, 6a and 6b.

XX SQ Sequence 7 AA;

Query Match 58.1%; Score 18; DB 18; Length 7;  
Best Local Similarity 57.1%; Pred. No. 7.8e+05;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMSTRES 7  
| | | | :  
Db 1 LLSPRGS 7

## RESULT 15

AAV05016

ID AAV05016 standard; peptide; 7 AA.

XX AC AAV05016;

XX DT 16-JUN-1999 (first entry)

XX DE Tumour antigen antibody light chain CDR2 clone F15.

XX KW Tumour antigen; antibody; CDR; complementarity determining region;  
KW binding molecule identification; tumour-specific binding polypeptide;  
KW cancer therapy; light chain.

XX OS Homo sapiens.

XX PN WO9906834-A2.

XX PD 11-FEB-1999.

XX PF 04-AUG-1998; 98WO-US16280.

XX PR 04-AUG-1997; 97US-0905825.

XX PR 04-AUG-1997; 97US-1112222.

XX PA (IXSY-) IXSYS INC.

XX PI Huse WD, Watkins JD, Wu H;

XX DR WPI; 1999-153951/13.

XX DR N-PSDB; AAX28193.

XX Identifying binding molecules for ligands, particularly tumour  
PT antigens - by selectively immobilising a population of binding  
PT molecules to a solid support and screening for binding to two or  
PT more ligands  
XX  
PS Claim 15; Page 57; 80pp; English.  
XX  
CC This sequence represents a light chain complementarity determining  
CC region (CDR) from a tumour antigen specific antibody.  
CC The invention relates to a method for identifying a binding molecule  
CC having selective affinity for a ligand comprising: (a) selectively  
CC immobilising a diverse population of binding molecules to a solid  
CC support; (b) simultaneously contacting the diverse population immobilised  
CC on the solid support with 2 or more ligands; and (c) determining at least  
CC one binding molecule which selectively binds to one or more of the  
CC ligands. The method allows for the rapid and efficient methods for the  
CC identification of binding molecules which exhibit selective affinity for  
CC one or more ligands of interest. They are used particularly for  
CC identifying tumour-specific binding polypeptides which can be used as  
CC targeting agents for cancer therapy that minimises impact on non-tumour  
XX tissues.  
XX  
SQ Sequence 7 AA;

Query Match 58.1%; Score 18; DB 20; Length 7;  
Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 STRAS 7  
|:|  
Db 3 SSRAS 7

Search completed: April 22, 2003, 13:21:55  
Job time : 35 secs

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:23:45 ; Search time 15 Seconds  
(without alignments)  
37.394 Million cell updates/sec

Title: US-09-674-716B-5

Perfect score: 31  
Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 18120

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pdb.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pdb.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pdb.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pdb.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pdb.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pdb.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pdb.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pdb.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pdb.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pdb.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pdb.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pdb.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	61.3	7	9	US-09-977-797A-30
2	18	58.1	7	9	US-09-977-797A-32
3	17	54.8	7	9	US-10-146-305-10
4	17	54.8	7	9	US-09-249-011A-18
5	17	54.8	7	9	US-10-144-644-11
6	17	54.8	7	10	US-09-828-708-43
7	17	54.8	7	10	US-09-828-708-46
8	16	51.6	7	10	US-09-947-137-8
9	15	48.4	7	9	US-10-078-458-11
10	15	48.4	6	10	US-09-800-433-8
11	15	48.4	6	12	US-10-014-774-19
12	15	48.4	7	9	US-09-995-973-53
13	15	48.4	7	9	US-10-116-255-58
14	15	48.4	7	9	US-09-913-238-62
15	15	48.4	7	10	US-09-850-351A-75
16	15	48.4	7	10	US-09-839-497A-17
17	15	48.4	7	10	US-09-828-708-45
18	15	48.4	7	10	US-09-828-708-47
19	15	48.4	7	10	US-09-828-708-48

20	15	48.4	7	10	US-09-828-708-49
21	14	45.2	3	10	US-09-982-172-82
22	14	45.2	5	10	US-09-800-433-7
23	14	45.2	5	10	US-09-976-165-63
24	14	45.2	6	9	US-09-858-852A-66
25	14	45.2	6	9	US-09-925-179-35
26	14	45.2	6	10	US-09-802-077-35
27	14	45.2	6	10	US-09-802-096-35
28	14	45.2	6	10	US-09-736-959A-22
29	14	45.2	6	10	US-09-990-762-66
30	14	45.2	7	9	US-09-977-797A-82
31	14	45.2	7	9	US-09-967-237-107
32	14	45.2	7	9	US-10-217-957-3
33	14	45.2	7	9	US-10-217-957-4
34	14	45.2	7	10	US-09-765-086-71
35	14	45.2	7	10	US-09-253-794-21
36	14	45.2	7	10	US-09-947-137-29
37	13	41.9	4	10	US-09-220-920-71
38	13	41.9	4	10	US-09-839-447A-78
39	13	41.9	4	10	US-09-782-980-102
40	13	41.9	4	10	US-09-405-499-3
41	13	41.9	4	10	US-09-872-349-1
42	13	41.9	4	10	US-09-982-172-87
43	13	41.9	4	10	US-09-751-100B-15
44	13	41.9	5	9	US-09-933-999A-30
45	13	41.9	5	10	US-09-287-849-36

## ALIGNMENTS

RESULT 1  
US-09-977-797A-30  
; Sequence 30, Application US/09977797A  
; Publication No. US20030044772A1  
; GENERAL INFORMATION:  
; APPLICANT: Watkins, Jeffrey D.  
; APPLICANT: Huse, William D.  
; APPLICANT: Wu, Herren  
; TITLE OF INVENTION: Methods for Identifying Ligand Specific Binding Molecules  
; FILE REFERENCE: AME-06805  
; CURRENT APPLICATION NUMBER: US/09/977,797A  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 09/129,026  
; PRIOR APPLICATION NUMBER: 08/905,825  
; PRIOR FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 136  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-797A-30

Query Match 61.3%; Score 19; DB 9; Length 7;  
Best Local Similarity 80.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 STRAS 7  
|||  
Db 3 STRAT 7

RESULT 2  
US-09-977-797A-32  
; Sequence 32, Application US/09977797A  
; Publication No. US20030044772A1  
; GENERAL INFORMATION:  
; APPLICANT: Watkins, Jeffrey D.  
; APPLICANT: Huse, William D.  
; APPLICANT: Wu, Herren  
; TITLE OF INVENTION: Methods for Identifying Ligand Specific Binding Molecules

```
; FILE REFERENCE: AME-06805
; CURRENT APPLICATION NUMBER: US/09/977,797A
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 09/129,026
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 08/905,825
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-797A-32

Query Match          58.1%; Score 18; DB 9; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 STRAS 7
   |||||
Db 3 SSRAS 7

RESULT 3
US-10-146-305-10
; Sequence 10, Application US/10146305
; Patent No. US20020173035A1
; GENERAL INFORMATION:
; APPLICANT: YUHAN CORPORATION
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV
; FILE REFERENCE: OVI7440
; CURRENT APPLICATION NUMBER: US/10/146,305
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: KR 10-2001-26634
; PRIOR FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-146-305-10

Query Match          54.8%; Score 17; DB 9; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRAS 7
   |||||
Db 3 STRES 7

RESULT 4
US-09-249-011A-18
; Sequence 19, Application US/09249011A
; Patent No. US20020176855A1
; GENERAL INFORMATION:
; APPLICANT: CO, MAN SUNG
; APPLICANT: VASQUEZ, MAXIMILIANO
; APPLICANT: CARRENO, BEATRIZ
; APPLICANT: CELNIKER, ABBIE CHERYL
; APPLICANT: COLLINS, MARY
; APPLICANT: GOLDMAN, SAMUEL
; APPLICANT: GRAY, GARY S.
; APPLICANT: KNIGHT, ANDREA
; APPLICANT: O'HARA, DENISE
; APPLICANT: RUP, BONITA
; APPLICANT: VELDMAN, GEERTUIDA M.
; TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH B7-2 AND METHODS
; TITLE OF INVENTION: OF TREATMENT THEREWITH
; FILE REFERENCE: 08702.0081-00000
```

```
; CURRENT APPLICATION NUMBER: US/09/249,011A
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CDR2 of humanized
; OTHER INFORMATION: murine anti-human B7-2 light chain
US-09-249-011A-18

Query Match          54.8%; Score 17; DB 9; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRAS 7
   |||||
Db 3 STRES 7

RESULT 5
US-10-144-644-11
; Sequence 11, Application US/10144644
; Publication No. US20030059429A1
; GENERAL INFORMATION:
; APPLICANT: Ames, Robert S.
; APPLICANT: Appelbaum, Edward R.
; APPLICANT: Chaiken, Irwin M.
; APPLICANT: Cook, Richard M.
; APPLICANT: Gross, Mitchell S.
; APPLICANT: Holmes, Stephen D.
; APPLICANT: McMillan, Lynette J.
; APPLICANT: Theisen, Timothy W.
; TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
; TITLE OF INVENTION: Treatment of IL5 Mediated Disorders
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp
; STREET: P.O. Box 1539-UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/144,644
; FILING DATE: 2002-08-08
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470110
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/467420
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363131
; FILING DATE: 23-DEC-1994
; ATTORNEY/AGENT INFORMATION: INFORMATION:
; NAME: Sutton, Jeffrey A.
; REGISTRATION NUMBER: 34,028
; REFERENCE/DOCKET NUMBER: P50282-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5024
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-10-144-644-11

Query Match      54.8%; Score 17; DB 9; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRAS 7
Db 3 STRES 7

RESULT 6
US-09-828-708-43
; Sequence 43, Application US/09828708
; Patent No. US20020146753A1
; GENERAL INFORMATION:
; APPLICANT: Burton, D.
; APPLICANT: Ditzel, H.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their partic
; FILE REFERENCE: 1361.005US1
; CURRENT APPLICATION NUMBER: US/09/828,708
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-708-43

Query Match      54.8%; Score 17; DB 10; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRAS 7
Db 3 STRES 7

RESULT 7
US-09-828-708-46
; Sequence 46, Application US/09828708
; Patent No. US20020146753A1
; GENERAL INFORMATION:
; APPLICANT: Burton, D.
; APPLICANT: Ditzel, H.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their partic
; FILE REFERENCE: 1361.005US1
; CURRENT APPLICATION NUMBER: US/09/828,708
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-708-46

Query Match      54.8%; Score 17; DB 10; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRAS 7
Db 3 STRES 7

RESULT 8
US-09-947-137-8
; Sequence 8, Application US/09947137
; Patent No. US20020137023A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Bruce F.
; APPLICANT: Samoilova, Tatiana
; TITLE OF INVENTION: Methods and Compositions for Targeting
; FILE REFERENCE: 5721-8
; CURRENT APPLICATION NUMBER: US/09/947,137
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 09/084,605
; PRIOR FILING DATE: 1998-05-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phage display library peptides
US-09-947-137-8

Query Match      51.6%; Score 16; DB 10; Length 7;
Best Local Similarity 60.0%; Pred. No. 2.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MSTRAS 6
Db 1 MSTQS 5

RESULT 9
US-10-078-458-11
; Sequence 11, Application US/10078458
; Publication No. US20020187934A1
; GENERAL INFORMATION:
; APPLICANT: LUKE, Richard William Arthur
; APPLICANT: COTTON, Ronald
; TITLE OF INVENTION: PEPTIDE DERIVATIVES
; FILE REFERENCE: 1991-174
; CURRENT APPLICATION NUMBER: US/10/078,458
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US/09/308,175
; PRIOR FILING DATE: 1999-05-17
; PRIOR APPLICATION NUMBER: PCT/GB97/03199
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: GB 9624562.6
; PRIOR FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-078-458-11

Query Match      48.4%; Score 15; DB 9; Length 4;
Best Local Similarity 75.0%; Pred. No. 2.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TRAS 7
Db 1 TRAA 4

RESULT 10
```

US-09-800-433-8  
; Sequence 8, Application US/09800433  
; Patent No. US20020106378A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Hare and Elliott  
; TITLE OF INVENTION: Materials and methods for intracellular transport and  
; TITLE OF INVENTION: their uses  
; FILE REFERENCE: 49408  
; CURRENT APPLICATION NUMBER: US/09/800,433  
; CURRENT FILING DATE: 2000-03-05  
; PRIOR APPLICATION NUMBER: 09/395,344  
; PRIOR FILING DATE: 1999-09-13  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: peptide  
US-09-800-433-8

Query Match 48.4%; Score 15; DB 10; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TRAS 7  
|||  
Db 2 TRAA 5

RESULT 11  
US-10-014-774-19  
; Sequence 19, Application US/10014774  
; Patent No. US2002009173A1  
; GENERAL INFORMATION:  
; APPLICANT: D'Azzo, Alessandra  
; APPLICANT: Bongiovanni, Antonella  
; APPLICANT: Nastasi, Tommaso  
; TITLE OF INVENTION: Protein Specific for Cardiac and Skeletal Muscle  
; FILE REFERENCE: 2427/1F509-US1  
; CURRENT APPLICATION NUMBER: US/10/014,774  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: PCT/US00/11900  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/131,814  
; PRIOR FILING DATE: 1999-04-29  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-014-774-19

Query Match 48.4%; Score 15; DB 12; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TRAS 7  
|||  
Db 2 TRAT 5

RESULT 12  
US-09-995-973-53  
; Sequence 53, Application US/09995973  
; Publication No. US20030024006A1  
; GENERAL INFORMATION:  
; APPLICANT: CHOO, Yen  
; APPLICANT: ULLMAN, Christopher G.  
; TITLE OF INVENTION: GENE SWITCHES  
; FILE REFERENCE: 8325-2003 / G7-US1

US-09-995-973-53  
; CURRENT APPLICATION NUMBER: US/09/995,973  
; CURRENT FILING DATE: 2002-03-19  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 53  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: zinc finger  
; OTHER INFORMATION: binding domain  
US-09-995-973-53

Query Match 48.4%; Score 15; DB 9; Length 7;  
Best Local Similarity 75.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTR 5  
|:  
Db 3 LVKTR 7

RESULT 13  
US-10-116-255-58  
; Sequence 58, Application US/10116255  
; Publication No. US20030036646A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: Uncoupling Protein Polynucleotides, Polypeptides, and  
; TITLE OF INVENTION: Antibodies  
; FILE REFERENCE: PT009P1  
; CURRENT APPLICATION NUMBER: US/10/116,255  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/685,897  
; PRIOR FILING DATE: 2000-10-11  
; PRIOR APPLICATION NUMBER: PCT/US00/09534  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: 60/128,701  
; PRIOR FILING DATE: 1999-04-09  
; PRIOR APPLICATION NUMBER: 60/142,821  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: 60/149,448  
; PRIOR FILING DATE: 1999-08-18  
; PRIOR APPLICATION NUMBER: 60/164,751  
; PRIOR FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 58  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-255-58

Query Match 48.4%; Score 15; DB 9; Length 7;  
Best Local Similarity 60.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LMSTR 5  
|:  
Db 3 LVKTR 7

RESULT 14  
US-09-913-238-62  
; Sequence 62, Application US/09913238  
; Publication No. US20030049251A1  
; GENERAL INFORMATION:  
; APPLICANT: Barbas, Carlos F.  
; APPLICANT: Steinberger, Peter  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR  
; TITLE OF INVENTION: INHIBITING CCR5-DEPENDENT INFECTION OF CELLS BY HIV-1  
; FILE REFERENCE: TSRI 728.1  
; CURRENT APPLICATION NUMBER: US/09/913,238



```

; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: PCT/EP 0/12419
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,653
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-913-238-62

```

```

Query Match 48.4%; Score 15; DB 9; Length 7;
Best Local Similarity 60.0%; Pred. No. 2.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 STRAS 7
Db 3 SSRSS 7

```

# RESULT 15

```

US-09-850-351A-75
; Sequence 75, Application US/09850351A
; Patent No. US20020100080A1
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Gerald S.
; SCHNEPF, H. Ernest
; Narva, Kenneth E.
; Stockhoff, Brian A.
; Schmeits, James
; Loewer, David
; Dullum, Charles Joseph
; Muller-Cohn, Judy
; Stamp, Lisa
; Morrill, George

```

```

TITLE OF INVENTION: No. US20020100080A1 Pesticidal Toxins and Nucleotide
Sequences Which Encode These Toxins

```

```

NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/850,351A
FILING DATE: 07-May-2001
CLASSIFICATION: <Unknown>

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/073,898
FILING DATE: 06-MAY-1998
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708CD1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 75:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-09-850-351A-75

```

```

Query Match 48.4%; Score 15; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 2.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LMST 4
Db 2 LLST 5

```

```

Search completed: April 22, 2003, 13:29:09
Job time : 16 secs

```



GenCore version 5.1.4 ps 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:21:20 ; Search time 28 seconds  
(without alignments)  
7.356 Million cell updates/sec

Title: US-09-674-716B-5

Perfect score: 31

Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 42808

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/6CTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	7	2	US-08-672-345C-23
2	31	100.0	7	2	US-08-672-345C-26
3	31	100.0	7	2	US-08-672-345C-29
4	31	100.0	7	4	US-09-214-095D-23
5	31	100.0	7	4	US-09-214-095D-26
6	31	100.0	7	4	US-09-214-095D-29
7	28	90.3	7	2	US-08-672-345C-20
8	28	90.3	7	2	US-09-214-095D-20
9	27	87.1	7	2	US-08-672-345C-80
10	27	87.1	7	4	US-09-214-095D-80
11	19	61.3	4	3	US-08-685-808-9
12	19	61.3	4	4	US-08-505-860C-9
13	18	58.1	7	1	US-08-353-400-31
14	17	54.8	7	1	US-08-264-093-25
15	17	54.8	7	1	US-08-467-420A-11
16	17	54.8	7	1	US-08-470-110A-11
17	17	54.8	7	1	US-08-360-125-20
18	17	54.8	7	1	US-08-667-789A-11
19	17	54.8	7	2	US-08-452-724A-34
20	17	54.8	7	2	US-08-450-578-20
21	17	54.8	7	2	US-08-940-371-11
22	17	54.8	7	2	US-09-017-628-20
23	17	54.8	7	3	US-09-014-880-20
24	17	54.8	7	3	US-09-100-409A-30
25	17	54.8	7	3	US-08-637-647-11
26	17	54.8	7	4	US-07-987-264-5
27	17	54.8	7	4	US-08-450-363-20

28	17	54.8	7	5	PCT-US93-08435-24	Sequence 24, Appl
29	17	54.8	7	5	PCT-US93-08435-37	Sequence 37, Appl
30	17	54.8	7	5	PCT-US93-08435-41	Sequence 41, Appl
31	17	54.8	7	5	PCT-US93-17082A-11	Sequence 11, Appl
32	16	51.6	4	1	US-07-776-257-3	Sequence 3, Appl
33	16	51.6	7	1	US-08-503-062-14	Sequence 14, Appl
34	16	51.6	7	2	US-08-340-283-90	Sequence 90, Appl
35	16	51.6	7	4	US-09-084-605B-8	Sequence 8, Appl
36	16	51.6	7	4	US-09-174-216-1	Sequence 1, Appl
37	16	51.6	7	5	PCT-US96-11495-14	Sequence 14, Appl
38	15	48.4	4	4	US-09-308-175A-11	Sequence 11, Appl
39	15	48.4	4	4	US-09-082-358B-47	Sequence 47, Appl
40	15	48.4	4	5	PCT-US93-01669-38	Sequence 38, Appl
41	15	48.4	5	1	US-08-064-111C-10	Sequence 10, Appl
42	15	48.4	6	2	US-08-729-152-7	Sequence 7, Appl
43	15	48.4	6	2	US-08-729-152-28	Sequence 28, Appl
44	15	48.4	6	3	US-09-012-126-8	Sequence 8, Appl
45	15	48.4	6	4	US-09-395-344-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-08-672-345C-23  
; Sequence 23, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-23

Query Match 100.0%; Score 31; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7  
| | | | |  
Db 1 LMSTRAS 7

RESULT 2

```
US-08-672-345C-26
; Sequence 26, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-26

Query Match 100.0%; Score 31; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 1 LMSTRAS 7

RESULT 3
US-08-672-345C-29
; Sequence 29, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-29

Query Match 100.0%; Score 31; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 1 LMSTRAS 7

RESULT 4
US-09-214-095D-23
; Sequence 23, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-23

Query Match 100.0%; Score 31; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 1 LMSTRAS 7

RESULT 5
US-09-214-095D-26
; Sequence 26, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-26

Query Match 100.0%; Score 31; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7
Db 1 LMSTRAS 7
```

Qy 1 LMSTRAS 7  
|||||  
Db 1 LMSTRAS 7

## RESULT 6

US-09-214-095D-29  
; Sequence 29, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 29  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-29

Query Match 100.0%; Score 31; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7  
|||||  
Db 1 LMSTRAS 7

## RESULT 7

US-08-672-345C-20  
; Sequence 20, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: /08/672,345C  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-20

Query Match 90.3%; Score 28; DB 2; Length 7;

Best Local Similarity 85.7%; Pred. No. 1.9e+05;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7  
|||||  
Db 1 LMSTRSS 7

## RESULT 8

US-09-214-095D-20  
; Sequence 20, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-20

Query Match 90.3%; Score 28; DB 4; Length 7;  
Best Local Similarity 85.7%; Pred. No. 1.9e+05;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7  
|||||  
Db 1 LMSTRSS 7

## RESULT 9

US-08-672-345C-80  
; Sequence 80, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 80:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-80

US-08-672-345C-80

Query Match 87.1%; Score 27; DB 2; Length 7;  
Best Local Similarity 85.7%; Pred. No. 1.9e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
| | | | |  
Db 1 LMSTRXS 7

RESULT 10

US-09-214-095D-80  
; Sequence 80, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 80  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Murinae gen.sp.  
; FEATURE:  
; NAME/KEY: CHAIN  
; LOCATION: (6), (6)  
; OTHER INFORMATION: X at position 6 represents any amino acid  
US-09-214-095D-80

Query Match 87.1%; Score 27; DB 4; Length 7;  
Best Local Similarity 85.7%; Pred. No. 1.9e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
| | | | |  
Db 1 LMSTRXS 7

RESULT 11

US-08-685-808-9  
; Sequence 9, Application US/08695808  
; Patent No. 6048715  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, CHARLES A., et al  
; TITLE OF INVENTION: SEPARATION AND CONCENTRATION SYSTEMS BASED  
; ON SOLUBLE OLIGOSACCHARIDE BINDING DOMAINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RAE-VENTER LAW GROUP  
; STREET: 260 Sheridan Ave., Ste. 440  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,808  
; FILING DATE: 24-JULY-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/505,860  
; FILING DATE: 24-JULY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rae-Venter, Barbara  
; REGISTRATION NUMBER: 32,750

; REFERENCE/DOCKET NUMBER: CBDT.017.01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 328-4400  
; TELEFAX: (650) 328-4477  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: no  
US-08-685-808-9

Query Match 61.3%; Score 19; DB 3; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MSTR 5  
| | | |  
Db 1 MSTR 4

RESULT 12

US-08-505-860C-9  
; Sequence 9, Application US/08505860C  
; Patent No. 6174700  
; GENERAL INFORMATION:  
; APPLICANT: HAYNES, CHARLES A., et al  
; TITLE OF INVENTION: SEPARATION AND CONCENTRATION SYSTEMS BASED  
; ON SOLUBLE OLIGOSACCHARIDE BINDING DOMAINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RAE-VENTER LAW GROUP  
; STREET: 260 Sheridan Ave., Ste. 440  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/505,860C  
; FILING DATE: 24-JULY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rae-Venter, Barbara  
; REGISTRATION NUMBER: 32,750  
; REFERENCE/DOCKET NUMBER: CBDT.017.00US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 328-4400  
; TELEFAX: (650) 328-4477  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: no  
US-08-505-860C-9

Query Match 61.3%; Score 19; DB 4; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MSTR 5  
| | | |  
Db 1 MSTR 4

RESULT 13

US-08-353-400-31  
; Sequence 31, Application US/08353400  
; Patent No. 5665357  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 37  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/353,400  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9324819.3  
; FILING DATE: 03-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9411089.7  
; FILING DATE: 03-JUN-1994  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-353-400-31

Query Match 58.1%; Score 18; DB 1; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 STRAS 7  
Db 3 STRTS 7

RESULT 14  
US-08-264-093-25  
; Sequence 25, Application US/08264093  
; Patent No. 5639863  
; GENERAL INFORMATION:  
; APPLICANT: Michael D. Dan  
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES SPECIFIC TO  
; TITLE OF INVENTION: CELL CYCLE-INDEPENDENT GLIOMA SURFACE  
; TITLE OF INVENTION: ANTIGEN  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ridout & Maybee  
; STREET: 2300 Richmond-Adelaide Centre  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 2J7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette - 3.5 inch, 1.4 Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: MS-DOS 6.00  
; SOFTWARE: ASCII Editor  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/264,093  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA: No. 5639863 applicable  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lake, James R.  
; REGISTRATION NUMBER: 31081  
; REFERENCE/DOCKET NUMBER: NOVOP/106A/7551  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 868-1482  
; TELEFAX: (416) 362-0823  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not applicable  
; TOPOLOGY: linear  
US-08-264-093-25

Query Match 54.8%; Score 17; DB 1; Length 7;  
Best Local Similarity 66.7%; Pred. No. 1.9e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MSTRAS 7  
Db 2 LSYRAS 7

RESULT 15  
US-08-467-420A-11  
; Sequence 11, Application US/08467420A  
; Patent No. 5683892  
; GENERAL INFORMATION:  
; APPLICANT: Ames, Robert S.  
; APPLICANT: Appelbaum, Edward R.  
; APPLICANT: Chaiken, Irwin M.  
; APPLICANT: Cook, Richard M.  
; APPLICANT: Gross, Mitchell S.  
; APPLICANT: Holmes, Stephen D.  
; APPLICANT: McMillan, Lynette J.  
; APPLICANT: Theisen, Timothy W.  
; TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in  
; TITLE OF INVENTION: Treatment of IL5 Mediated Disorders  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corp./Corporate  
; ADDRESSEE: Intellectual Property  
; STREET: P. O. Box 1539-UW2220  
; CITY: King of Prussia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,420A  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/363131  
; FILING DATE: 23-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sutton, Jeffrey A.  
; REGISTRATION NUMBER: 34,028  
; REFERENCE/DOCKET NUMBER: P50282  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610 270-5024  
; TELEFAX: 610 270-5090  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-467-420A-11

Query Match 54.8%; Score 17; DB 1; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.9e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 STRAS 7  
 |||||  
 Db 3 STRES 7

Search completed: April 22, 2003, 13:25:06  
 Job time : 29 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:28:56 ; Search time 14 Seconds  
(without alignments)  
61.801 Million cell updates/sec

Title: US-09-674-716b-7

Perfect score: 48

Sequence: 1 QQLVEYPPT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 789

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	43.8	9	S66635	alpha-2-macroglobu
2	19	39.6	5	E60274	major protein anti
3	16	33.3	3	A43391	TRH-like tripeptid
4	14	29.2	4	A32039	tyrosine-melanocyt
5	14	29.2	4	PT0240	Ig heavy chain CRD
6	14	29.2	5	PQ0009	angiotensin-conver
7	14	29.2	6	JN0861	peptidyl-dipeptida
8	14	29.2	7	A61324	dermorphin - Ronde
9	14	29.2	7	S36662	dermorphin (Lys-7)
10	14	29.2	7	S21230	dermorphin (Trp-4,
11	14	29.2	7	S46620	aggreccan - bovine
12	14	29.2	8	S66646	cardioacceleratory
13	14	29.2	9	A61620	locustamyotropin I
14	14	29.2	9	B49712	ATP-binding protei
15	13	27.1	5	JS0319	subesophageal gang
16	13	27.1	6	A61049	halo-toxin - pseud
17	13	27.1	6	IS1317	bHLH transcription
18	13	27.1	7	E33932	Ig mu chain D regi
19	13	27.1	7	PX0008	glucuronosyltransf
20	13	27.1	7	I56695	hypothetical L2 pr
21	13	27.1	8	S16324	hypothetical prote
22	13	27.1	9	S15850	vitamin D3 26-mono
23	13	27.1	9	A61358	bradykinin-like pe
24	13	27.1	9	A61057	Thr-6 bradykinin -
25	13	27.1	9	A26744	bradykinin-like pe
26	13	27.1	9	A61363	bradykinin - commo
27	13	27.1	9	A60579	bradykinin-like pe
28	13	27.1	9	A26363	cardioactive pepti
29	13	27.1	9	S39766	cardioactive pepti

30	13	27.1	9	2	S27233	cardioactive pepti
31	13	27.1	9	2	S65433	bradykinin - horn
32	13	27.1	9	2	PT0080	60K Ca binding pro
33	13	27.1	9	2	PT0285	Ig heavy chain CRD
34	13	27.1	9	2	S10784	enamelin I - bovin
35	13	27.1	9	2	I46023	growth hormone rec
36	13	27.1	9	2	S39767	cardioactive pepti
37	13	27.1	9	2	A43065	hydroxyproline-3-b
38	12	25.0	5	2	B37325	pap fibribrial regul
39	12	25.0	5	3	JT0870	phytoalbumin albu
40	12	25.0	7	2	S25266	phile protein - bac
41	12	25.0	7	2	A38081	amine oxidase (cop
42	12	25.0	8	2	PL0184	capsid protein VP-
43	12	25.0	8	2	D47393	neuropeptide calla
44	12	25.0	8	2	E47393	neuropeptide calla
45	12	25.0	8	2	B45800	serum albumin - do

## ALIGNMENTS

### RESULT 1

S66635

alpha-2-macroglobulin isoform 1 - bovine (fragment)

C;Species: Bos primigenius indicus (zebu cattle)

C;Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999

C;Accession: S66635

R;Dolmer, K.; Jenner, L.B.; Jacobsen, L.; Andersen, G.R.; Koch, T.J.; Thirup, S.; Sottr

FEBS Lett. 372, 93-95, 1995

A;Title: Crystallisation and preliminary X-ray analysis of the receptor-binding domain o

A;Reference number: S66634; MUID:96032553; PMID:7556651

A;Accession: S66635

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-9 <DOL>

Query Match 43.8%; Score 21; DB 2; Length 9;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EYPF 8

Db 3 EYPF 6

### RESULT 2

E60274

major protein antigen MPT63 - Mycobacterium tuberculosis (fragment)

C;Species: Mycobacterium tuberculosis

C;Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 30-Sep-1993

C;Accession: E60274

R;Negai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.

Infect. Immun. 59, 372-382, 1991

A;Title: Isolation and partial characterization of major protein antigens in the culture

A;Reference number: A60274; MUID:91099989; PMID:1898899

A;Accession: E60274

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-5 <NAG>

Query Match 39.6%; Score 19; DB 2; Length 5;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 YPPT 9

Db 2 YPIT 5

### RESULT 3

A43391

TRH-like tripeptide - alfalfa

C;Species: Medicago sativa (alfalfa)

C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C/Accession: A43391

R/Lackey, D.B.

J. Biol. Chem. 267, 17508-17511, 1992

A/Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroGlu-  
A/Reference number: A43391; MUID:92388092; PMID:1517203

A/Accession: A43391

A/Molecule type: protein

A/Residues: 1-3 <LAC>

C/Keywords: amidated carboxyl end; pyroglutamic acid

F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 33.3%; Score 16; DB 3; Length 3;

Best Local Similarity 66.7%; Pred. No. 2.8e+05;

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YP 7

Db 1 QVP 3

#### RESULT 4

A32039

tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 31-Jul-1989 #sequence\_revision 31-Jul-1989 #text\_change 18-Aug-2000

C/Accession: A32039

R/Horvath, A.; Kastin, A.J.

J. Biol. Chem. 264, 2175-2179, 1989

A/Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting factor

A/Reference number: A32039; MUID:89123285; PMID:2563371

A/Accession: A32039

A/Molecule type: protein

A/Residues: 1-4 <HOR>

A/Experimental source: brain

C/Superfamily: unassigned animal peptides

C/Keywords: amidated carboxyl end

F1/4/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match

Best Local Similarity 29.2%; Score 14; DB 2; Length 4;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YP 7

Db 1 YP 2

#### RESULT 5

PT0240

Ig heavy chain CRD3 region (clone 2-100B) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C/Accession: PT0240

R/Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A/Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J

A/Reference number: PT0222; MUID:91108337; PMID:1899102

A/Accession: PT0240

A/Molecule type: DNA

A/Residues: 1-4 <YAM>

A/Experimental source: B lymphocyte

C/Keywords: heterotetramer; immunoglobulin

Query Match

Best Local Similarity 29.2%; Score 14; DB 2; Length 4;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YP 7

Db 1 YP 2

#### RESULT 6

PQ0009

angiotensin-converting enzyme inhibitor (FLP-2) - common fig

N/Alternate names: ficus latex peptide 2

C/Species: Ficus carica (common fig)

C/Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 08-Dec-1995

C/Accession: PQ0009

R/Maryama, S.; Miyoshi, S.; Tanaka, H.

Agric. Biol. Chem. 53, 2763-2767, 1989

A/Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.

A/Reference number: PQ0008

A/Accession: PQ0009

A/Molecule type: protein

A/Residues: 1-5 <MAR>

A/Experimental source: latex

C/Keywords: angiotensin-converting enzyme inhibitor

Query Match

Best Local Similarity 29.2%; Score 14; DB 2; Length 5;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YP 7

Db 2 YP 3

#### RESULT 7

JN0861

peptidyl-dipeptidase A inhibitory peptide C111 - striped bonito

C/Species: Sarda orientalis (striped bonito)

C/Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 07-May-1999

C/Accession: JN0861

R/Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.

Biosci. Biotechnol. Biochem. 57, 1743-1744, 1993

A/Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory pe

A/Reference number: JN0859; MUID:94080036; PMID:7764272

A/Accession: JN0861

A/Molecule type: protein

A/Residues: 1-6 <MAR>

A/Experimental source: liver

C/Comment: The carboxyl end is essential for the protein's expression of angiotensin I-c

C/Superfamily: bradykinin-potentiating peptide

C/Keywords: angiotensin-converting enzyme inhibitor

Query Match

Best Local Similarity 29.2%; Score 14; DB 2; Length 6;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YP 7

Db 3 YP 4

#### RESULT 8

A61324

dermorphin - Rohde's leaf frog

C/Species: Phyllomedusa rohdei (Rohde's leaf frog)

C/Date: 17-Jul-1994 #sequence\_revision 05-Aug-1994 #text\_change 07-May-1999

C/Accession: A61324

R/Montecucchi, P.C.; De Castiglione, R.; Erspamer, V.

Int. J. Pept. Protein Res. 17, 316-321, 1981

A/Title: Identification of dermorphin and Hyp(6)-dermorphin in skin extracts of the Bra

A/Reference number: A61324; MUID:82029915; PMID:7287302

A/Accession: A61324

A/Molecule type: protein

A/Residues: 1-7 <MON>

C/Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

C/Keywords: amidated carboxyl end; D-amino acid; hydroxyproline; neuropeptide; skin

F2/Modified site: D-alanine (Ala) #status experimental

F6/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

F7/Modified site: amidated carboxyl end (Ser) #status experimental

Query Match 29.2%; Score 14; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YP 7  
 ||  
 Db 5 YP 6

RESULT 9  
 S36662  
 dermorphin (luis-7) [validated] - two-colored leaf frog  
 C:Species: Phyllomedusa bicolor (two-colored leaf frog)  
 C>Date: 10-Dec-1993 #sequence\_revision 13-Mar-1997 #text\_change 18-Aug-2000  
 C:Accession: S36662  
 R:Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil, G.  
 FEBS Lett. 302, 151-154, 1992  
 A:Title: Identification and characterization of two dermorphins from skin extracts of the  
 A:Reference number: S21152; MUID:92339502; PMID:1633846  
 A:Accession: S36662  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-7 <MIG>  
 C:Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

Query Match 29.2%; Score 14; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YP 7  
 ||  
 Db 5 YP 6

RESULT 10  
 S21230  
 dermorphin (Tnp-4, Asn-7) [validated] - two-colored leaf frog (fragment)  
 C:Species: Phyllomedusa bicolor (two-colored leaf frog)  
 C>Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 18-Aug-2000  
 C:Accession: S21230  
 R:Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil, G.  
 FEBS Lett. 302, 151-154, 1992  
 A:Title: Identification and characterization of two dermorphins from skin extracts of the  
 A:Reference number: S21152; MUID:92339502; PMID:1633846  
 A:Accession: S21230  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-7 <MIG>  
 C:Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

Query Match 29.2%; Score 14; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YP 7  
 ||  
 Db 5 YP 6

RESULT 11  
 S42620  
 aggrecan - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 19-Mar-1997 #sequence\_revision 24-Jul-1998 #text\_change 24-Jul-1998  
 C:Accession: S42620  
 R:Vogel, K.G.; Sandy, J.D.; Pogany, G.; Robbins, J.R.  
 Matrix Biol. 14, 171-179, 1994  
 A:Title: Aggrecan in bovine tendon.  
 A:Reference number: S42620; MUID:94340214; PMID:7520336  
 A:Accession: S42620  
 A:Molecule type: protein  
 A:Residues: 1-7 <VOG>  
 A:Experimental source: flexor tendon

C:Keywords: cartilage

Query Match 29.2%; Score 14; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YP 7  
 ||  
 Db 1 YP 2

RESULT 12  
 S66646  
 cardioacceleratory protein 2b - tobacco hornworm  
 C:Species: Manduca sexta (tobacco hornworm)  
 C>Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
 C:Accession: S66646  
 R:Huesmann, G.R.; Cheung, C.C.; Loi, P.K.; Lee, T.D.; Swiderek, K.M.; Tublitz, N.J.  
 FEBS Lett. 371, 311-314, 1995  
 A:Title: Amino acid sequence of CAP(2b), an insect cardioacceleratory peptide from the t  
 A:Reference number: S66646; MUID:96013159; PMID:7556618  
 A:Accession: S66646  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-8 <HUE>

Query Match 29.2%; Score 14; DB 2; Length 8;  
 Best Local Similarity 33.3%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QLVYYP 7  
 :||  
 Db 1 ELYAFP 6

RESULT 13  
 A61620  
 locustamytropin III - migratory locust  
 C:Species: Locusta migratoria (migratory locust)  
 C>Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 11-Jul-1997  
 C:Accession: A61620  
 R:Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; Kochansky, J.P.; De Loof, A.  
 Insect Biochem. Mol. Biol. 22, 447-452, 1992  
 A:Title: Isolation, identification and synthesis of locustamytropin III and IV, two add  
 A:Reference number: A61620  
 A:Accession: A61620  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9 <SCH>  
 C:Keywords: amidated carboxyl end; neuropeptide  
 F:9/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 29.2%; Score 14; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYPF 8  
 :||  
 Db 2 QQPF 5

RESULT 14  
 B49712  
 ATP-binding protein p46 - dog (fragment)  
 C:Species: Canis lupus familiaris (dog)  
 C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999  
 C:Accession: B49712  
 R:Nigam, S.K.; Goldberg, A.L.; Ho, S.; Rohde, M.F.; Bush, K.T.; Sherman, M.Y.  
 J. Biol. Chem. 269, 1744-1749, 1994  
 A:Title: A set of endoplasmic reticulum proteins possessing properties of molecular chap  
 A:Reference number: A49712; MUID:94124514; PMID:8294423  
 A:Accession: B49712  
 A:Status: preliminary

A:Molecule type: protein  
A:Residues: 1-9 <NIG>  
C:Keywords: ATP; endoplasmic reticulum; molecular chaperone

Query Match 29.2%; Score 14; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YP 7  
Db 5 YP 6

## RESULT 15

JS0319  
subesophageal ganglion pentapeptide - house cricket  
C:Species: Acheta domesticus (house cricket)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Jun-2000  
C:Accession: JS0319  
R:Wicker, C.; Wicker, C.  
Comp. Biochem. Physiol. C 88, 185-187, 1987  
A>Title: Isolation and structure of a peptide isolated from the subesophageal ganglion  
A:Reference number: JS0319  
A:Accession: JS0319  
A:Molecule type: protein  
A:Residues: 1-5 <NIG>

Query Match 27.1%; Score 13; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PF 8  
Db 4 PF 5

Search completed: April 22, 2003, 13:31:56  
Job time : 14 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:25:11 ; Search time 24 Seconds  
(without alignments)  
15.554 Million cell updates/sec

Title: US-09-674-716B-7  
Perfect score: 48  
Sequence: 1 QQLVEYPT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 231

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	15	31.2	8	ANG2 BOTJA	Q10582 bothrops ja
2	15	31.2	8	UPAA HUMAN	P30096 homo sapien
3	15	31.2	9	FLA2 TREHY	P80159 treponema h
4	14	29.2	8	AL17 CARMA	P81820 carcinus ma
5	14	29.2	9	AL11 CARMA	P81814 carcinus ma
6	14	29.2	9	LMT3 LOEMI	P41489 locusta mig
7	14	29.2	9	UPA3 HUMAN	P30089 homo sapien
8	14	29.2	9	UPA7 HUMAN	P30093 homo sapien
9	13	27.1	5	SUGA ACHDO	P19991 acheta dome
10	13	27.1	7	FAR1 HELTI	P41871 helisoma tr
11	13	27.1	7	UN06 PINPS	P81675 pinus pinas
12	13	27.1	8	PPK2 PERAM	P82692 periplaneta
13	13	27.1	8	PPK3 PERAM	P82618 periplaneta
14	13	27.1	9	CCAP CARMA	P38556 carcinus ma
15	13	27.1	9	UPA2 HUMAN	P40929 homo sapien
16	12	25.0	5	AL14 CARMA	P81817 carcinus ma
17	12	25.0	5	PSK DAUCA	P58261 daucus caro
18	12	25.0	6	ASP2 LACSN	P82655 lactobacill
19	12	25.0	7	AL12 CARMA	P81805 carcinus ma
20	12	25.0	7	AL13 CARMA	P81806 carcinus ma
21	12	25.0	7	AL14 CARMA	P81807 carcinus ma
22	12	25.0	7	AL15 CARMA	P81808 carcinus ma
23	12	25.0	7	AL17 CYDPO	P82158 cydia pomon
24	12	25.0	7	FAR2 ASCSU	P31890 ascaris suu
25	12	25.0	8	AL12 CARMA	P81815 carcinus ma
26	12	25.0	8	AL13 CARMA	P81818 carcinus ma
27	12	25.0	8	AL18 CARMA	P81821 carcinus ma
28	12	25.0	8	ALL3 CYDPO	P82154 cydia pomon
29	12	25.0	8	ALL4 CALVO	P41840 calliphora
30	12	25.0	8	ALL4 CYDPO	P82155 cydia pomon
31	12	25.0	8	ALL5 CALVO	P41841 calliphora
32	12	25.0	8	ALL5 CYDPO	P82156 cydia pomon
33	12	25.0	8	ALL7 CARMA	P81809 carcinus ma

P81811 carcinus ma  
P81812 carcinus ma  
P22396 locusta mig  
P81813 carcinus ma  
P80975 thunnus obe  
P04277 homo sapien  
P82072 litoria rub  
P01151 sus scrofa  
P82071 litoria rub  
P82073 litoria rub  
P25418 libellula a  
P14595 tabanus atr

34 12 25.0 8 1 ALL8 CARMA  
35 12 25.0 8 1 ALL9 CARMA  
36 12 25.0 8 1 LMT2 LOEMI  
37 12 25.0 9 1 ALL10 CARMA  
38 12 25.0 9 1 COXE THUOB  
39 12 25.0 9 1 NEUX HUMAN  
40 11.5 24.0 5 1 RE31 LITRU  
41 11 22.9 3 1 THYL PIG  
42 11 22.9 5 1 RE21 LITRU  
43 11 22.9 5 1 RE32 LITRU  
44 11 22.9 8 1 AKH LIBAU  
45 11 22.9 8 1 AKH TABAT

## ALIGNMENTS

RESULT 1  
ANG2 BOTJA STANDARD; PRT; 8 AA.  
AC Q10582;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Angiotensin-like peptide II (Fragment).  
OS Bothrops jararaca (Jararaca).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
OC Viperidae; Crotalinae; Bothrops.  
OX NCBI\_TaxID=8724;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Plasma;  
RX MEDLINE=96208932; PubMed=8829801;  
RA Borgheresi R.A.M.B., Dalle Lucca J., Carmona E., Picarelli Z.P.;  
RT "Isolation and identification of angiotensin-like peptides from the  
plasma of the snake Bothrops jararaca.";  
RL Comp. Biochem. Physiol. 113B:467-473(1996).  
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
DR InterPro: IPR000215; Serpin.  
DR PROSITE; PS00284; SERPIN; PARTIAL.  
KW Vasoconstrictor; Plasma; Serpin.  
FT NON TER 8  
SQ SEQUENCE 8 AA; 1046 MW; DDD761E04B42D40A CRC64;

Query Match 31.2%; Score 15; DB 1; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.1e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 YPF 8  
: : :  
DB 6 HPF 8

RESULT 2  
UPAA HUMAN STANDARD; PRT; 8 AA.  
AC P30096;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Unknown protein from 2D-page of plasma (Spot 36) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Plasma;  
RX MEDLINE=93092937; PubMed=1459097;  
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
Hochstrasser D.F.;

RT "plasma protein map: an update by microsequencing.";  
 RL Electrophoresis 13:707-714(1992).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 7, ITS MW IS: 12 kDa.  
 DR SWISS-2DPAGE; P30096; HUMAN.  
 FT NON TER 1  
 FT VARIANT 5  
 FT 5 F -> P.  
 FT 8 /FTID=VAR\_000004.  
 FT NON TER 8  
 FT SEQUENCE 8 AA; 909 MW; 86677B59D1A72042 CRC64;  
 Query Match 31.2%; Score 15; DB 1; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 1.1e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 LVEYFP 7  
 Db 3 LTFYFP 7  
 RESULT 3  
 ID FLA2\_TREHY STANDARD; PRT; 9 AA.  
 AC P80159;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Flagellar filament outer layer protein flaa2 (35 kDa sheath protein)  
 DE (Fragment).  
 GN FLAA2.  
 OS Treponema hyodysenteriae (Serpulina hyodysenteriae).  
 OC Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.  
 OX NCBI\_TaxID=159;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=CS;  
 RX MEDLINE=93139764; PubMed=1487733;  
 RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,  
 RA van der Zeijst B.A.M., Kusters J.G.;  
 RT "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are  
 RT composed of two sheath proteins and three core proteins.";  
 RL J. Gen. Microbiol. 138:2697-2706(1992).  
 CC -!- SUBUNIT: COMPONENT OF THE OUTER LAYER OF THE FLAGELLA.  
 CC SHEATH PROTEINS. FLAA1 (44 kDa) AND FLAA2 (35 kDa) AROUND A CORE  
 CC THAT CONTAINS THREE PROTEINS FLAB1 (37 kDa), FLAB2 (34 kDa) AND  
 CC FLAB3 (32 kDa).  
 CC -!- SUBCELLULAR LOCATION: Periplasmic flagellum.  
 KW Flagella; Periplasmic.  
 FT UNSURE 2  
 FT UNSURE 8  
 FT NON TER 9  
 FT SEQUENCE 9 AA; 1129 MW; 855A19C68B4772D1 CRC64;  
 Query Match 31.2%; Score 15; DB 1; Length 9;  
 Best Local Similarity 42.9%; Pred. No. 1.1e+05;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 QLVEYFP 8  
 Db 1 ETVPYMF 7  
 RESULT 4  
 ID AL17\_CARMA STANDARD; PRT; 8 AA.  
 AC P81820;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Carcinus maenas 17.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;

OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 OC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD RES 8  
 FT SEQUENCE 8 AA; 858 MW; C82879D5A846D865 CRC64;  
 Query Match 29.2%; Score 14; DB 1; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 EYFP 8  
 Db 3 QYSF 6  
 RESULT 5  
 ID AL11\_CARMA STANDARD; PRT; 9 AA.  
 AC P81814;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Carcinustatin 11.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 OC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD RES 9  
 FT SEQUENCE 9 AA; 927 MW; 832D79DCB46D861 CRC64;  
 Query Match 29.2%; Score 14; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 EYFP 8  
 Db 4 QYAF 7  
 RESULT 6  
 ID LMT3\_LOCMI STANDARD; PRT; 9 AA.  
 AC P41489;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Locustamyotropin 3 (LOW-MT-3).  
 OS Locusta migratoria (Migratory locust).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
 OC Acridoidea; Acridoidea; Acrididae; Oedipodinae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Brain;  
 RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., Kochansky J.P.,  
 RA de Loof A.;  
 RT "Isolation, identification and synthesis of locustamyotropin III and  
 RT IV, two additional neuropeptides of Locusta migratoria: members of the  
 RT locustamyotropin peptide family.";  
 RL Insect Biochem. Mol. Biol. 22:447-452(1992).  
 CC -|- FUNCTION: POTENT MEDIATOR OF VISCERAL MUSCLE CONTRACTILE ACTIVITY  
 CC (MYOTROPIC ACTIVITY).  
 CC -|- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.  
 DR InterPro; IPR001484; Pyrokinin.  
 DR PROSITE; PS00539; PYROKININ; 1.  
 KW Neuropeptide; Amidation; Pyrokinin.  
 FT MOD RES 9  
 FT SEQUENCE 9 AA; 1140 MW; DSAE172C9D776C6 CRC64;  
 SQ SEQUENCE 9 AA; 1140 MW; DSAE172C9D776C6 CRC64;

Query Match 29.2%; Score 14; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYPF 8  
 : ||  
 Db 2 QOPF 5

## RESULT 7

UPA3\_HUMAN  
 ID UPA3\_HUMAN STANDARD; PRT; 9 AA.  
 AC P30089;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 2D-page of plasma (Spot 11) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing.";  
 RT Electrophoresis 13:707-714(1992).  
 RL MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC -|- PROTEIN IS: 4.6, ITS MW IS: 46 kDa.  
 CC SWISS-2DPAGE; P30089; HUMAN.  
 DR PIR; JS0319;  
 FT NON\_TER 1  
 FT NON\_TER 9  
 FT SEQUENCE 9 AA; 1056 MW; 26F2B1BAF769C737 CRC64;

Query Match 29.2%; Score 14; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 YPFT 9  
 : ||  
 Db 4 FPXT 7

## RESULT 8

UPA7\_HUMAN  
 ID UPA7\_HUMAN STANDARD; PRT; 9 AA.  
 AC P30093;  
 DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 2D-page of plasma (Spot 18) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing.";  
 RT Electrophoresis 13:707-714(1992).  
 RL MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC -|- PROTEIN IS: 5.05, ITS MW IS: 37 kDa.  
 CC SWISS-2DPAGE; P30093; HUMAN.  
 DR PIR; JS0319;  
 FT NON\_TER 1  
 FT NON\_TER 5  
 FT NON\_TER 9  
 FT SEQUENCE 9 AA; 1042 MW; 5C14477AEB0772C7 CRC64;

Query Match 29.2%; Score 14; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYPF 8  
 : ||  
 Db 5 EXPY 8

## RESULT 9

SUGA\_ACHDO  
 ID SUGA\_ACHDO STANDARD; PRT; 5 AA.  
 AC P19991;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Suboesophageal ganglion pentapeptide.  
 OS Acheta domestica (House cricket).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera;  
 OC Gryllidae; Gryllinae; Acheta.  
 OX NCBI\_TaxID=6997;  
 RN [1]  
 RP SEQUENCE.  
 RA Wicker C., Wicker C.;  
 RT "Isolation and structure of a peptide isolated from the  
 RT suboesophageal ganglion of Acheta domestica (orthoptera).";  
 RL Comp. Biochem. Physiol. 88C:185-187(1987).  
 CC -|- SUBCELLULAR LOCATION: MAIN PEPTIDE FROM THE SUBESOPHAGEAL  
 CC GANGLIA.  
 DR PIR; JS0319;  
 FT NON\_TER 1  
 FT NON\_TER 9  
 FT SEQUENCE 5 AA; 476 MW; 69D76DDDDDB00000 CRC64;

Query Match 27.1%; Score 13; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PF 8  
 : ||  
 Db 4 PF 5

## RESULT 10

FARI\_HEITI  
 ID FARI\_HEITI STANDARD; PRT; 7 AA.  
 AC P41871;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)

```

DE FMRamide-like neuropeptide GDFFLRP-amide.
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Basommatophora;
OC Planorbidae; Helisoma.
OX NCBI_TaxID=27815;
RN [1]
RP SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRamide-related peptides from the kidney of the snail, Helisoma
RL trivolvis.";
RT Peptides 15:31-36(1994).
CC -!- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING
CC -!- THE KIDNEY, MANTLE AND SKIN.
CC -!- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.
CC -!- SIMILARITY: BELONGS TO THE FMRP (FMRPAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;

Query Match 27.1%; Score 13; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PF 8
DB 3 PF 4

RESULT 11
UN06 PINPS STANDARD; PRT; 7 AA.
AC P81675;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Unknown protein from 2D-page of needles (N141) (Fragment).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC -!- PROTEIN IS: 6.6, ITS MW IS: 25 kDa.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 7
SQ SEQUENCE 7 AA; 823 MW; 69D76724486B5740 CRC64;

Query Match 27.1%; Score 13; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PF 8
DB 6 PF 7

RESULT 12
PPK2 PERAM STANDARD; PRT; 8 AA.
AC P82632;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE "Differential distribution of pyrokinin-isoforms in cerebral and
DE abdominal neurohemal organs of the American cockroach.";
INsect Biochem. Mol. Biol. 29:139-144(1999).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.";
J. Comp. Neurol. 419:352-363(2000).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -!- TISSUE SPECIFICITY: CORPORA CARDIACA.

DE FMRamide-like neuropeptide GDFFLRP-amide.
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Basommatophora;
OC Planorbidae; Helisoma.
OX NCBI_TaxID=27815;
RN [1]
RP SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRamide-related peptides from the kidney of the snail, Helisoma
RL trivolvis.";
RT Peptides 15:31-36(1994).
CC -!- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING
CC -!- THE KIDNEY, MANTLE AND SKIN.
CC -!- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.
CC -!- SIMILARITY: BELONGS TO THE FMRP (FMRPAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;

Query Match 27.1%; Score 13; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PF 8
DB 3 PF 4

RESULT 13
PPK3 PERAM STANDARD; PRT; 8 AA.
AC P82618;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE "Differential distribution of pyrokinin-isoforms in cerebral and
DE abdominal neurohemal organs of the American cockroach.";
INsect Biochem. Mol. Biol. 29:139-144(1999).
RN [1]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.";
J. Comp. Neurol. 419:352-363(2000).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -!- TISSUE SPECIFICITY: CORPORA CARDIACA.

```

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pyrokinin-2 (Pea-PK-2) (FXPRL-amide).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattodea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=97353923; PubMed=9210163;
RA Predel R., Kellner R., Kaufmann R., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of two pyrokinins from the
RT retrocerebral complex of the American cockroach.";
RL Peptides 18:473-478(1997).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.";
J. Comp. Neurol. 419:352-363(2000).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -!- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC -!- MASS SPECTROMETRY: MW=883; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR InterPro; IPR001484; Pyrokinin.
DR PROSITE; PS00539; PYROKININ; FALSE_NEG.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 8
FT MOD_RES 8
SQ SEQUENCE 8 AA; 884 MW; C834176DD9D77775 CRC64;

Query Match 27.1%; Score 13; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PF 8
DB 3 PF 4

RESULT 13
PPK3 PERAM STANDARD; PRT; 8 AA.
AC P82618;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE "Differential distribution of pyrokinin-isoforms in cerebral and
DE abdominal neurohemal organs of the American cockroach.";
INsect Biochem. Mol. Biol. 29:139-144(1999).
RN [1]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.";
J. Comp. Neurol. 419:352-363(2000).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -!- TISSUE SPECIFICITY: CORPORA CARDIACA.

```



CC -1- MASS SPECTROMETRY: MW=996.5; METHOD=WALDI.  
 CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.  
 KW Neuropeptide; Amidation; Pyrokinin.  
 FT MOD RES 8 8  
 SQ SEQUENCE 8 AA; 997 MW; 0B34177409D772C7 CRC64;

Query Match 27.1%; Score 13; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 PF 8  
 ||  
 Db 3 PF 4

RESULT 14  
 CCAP\_CARMA STANDARD; PRT; 9 AA.  
 AC P38556;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Cardioactive peptide (CCAP).  
 OS Carcinus maenas (Common shore crab) (Green crab),  
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm),  
 OS Tenebrio molitor (Yellow mealworm), and  
 OS Spodoptera eridania (Southern armyworm).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 OC Brachyura; Eubrachyura; Portunioidea; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759, 7130, 7067, 37547;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=C.maenas; TISSUE=Pericardial organs;  
 RA Stangier J., Hilbich C., Beyreuther K., Keller R.;  
 RT "Unusual cardioactive peptide (CCAP) from pericardial organs of the  
 RT shore crab Carcinus maenas.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:575-579 (1987).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=M.sexta;  
 RX MEDLINE=93050243; PubMed=1426284;  
 RA Cheung C.C., Loi P.K., Sylwester A.W., Lee T.D., Tublitz N.J.;  
 RT "Primary structure of a cardioactive neuropeptide from the tobacco  
 RT hawkmoth, Manduca sexta.";  
 RL FEBS Lett. 313:165-168 (1992).  
 RN [3]  
 RP SEQUENCE.  
 RC SPECIES=T.molitor, and S.eridania; TISSUE=Head;  
 RX MEDLINE=94176032; PubMed=8129851;  
 RA Furuya K., Liao S., Reynolds S.E., Ota R.B., Hackett M.,  
 RA Schooley D.A.;  
 RT "Isolation and identification of a cardioactive peptide from Tenebrio  
 RT molitor and Spodoptera eridania.";  
 RL Biol. Chem. Hoppe-Seyler 374:1065-1074 (1993).  
 CC -1- FUNCTION: THE EFFECT OF CCAP IS BOTH INO- AND CHRONOTROPIC.  
 CC -1- TISSUE SPECIFICITY: STORED IN PERICARDIAL ORGANS AND RELEASED  
 CC INTO THE HEMOLYMPH.  
 DR PIR; A26363; A26363.  
 DR PIR; S27233; S27233.  
 KW Neuropeptide; Amidation.  
 FT DISULFID 3 9  
 FT MOD RES 9 9  
 SQ SEQUENCE 9 AA; 959 MW; C5A861A9CDD44EB9 CRC64;

Query Match 27.1%; Score 13; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 PF 8  
 ||  
 Db 1 PF 2

RESULT 15  
 UHA2\_HUMAN STANDARD; PRT; 9 AA.  
 ID UHA2\_HUMAN  
 AC P40929;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 2D-page of heart (Spot 5603) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 RT 1994.";  
 RL Electrophoresis 15:1459-1465 (1994).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6.0, ITS MW IS: 55.3 kDa.  
 FT NON TER 9 9  
 SQ SEQUENCE 9 AA; 1104 MW; 8874B1BB5B01B2CA CRC64;

Query Match 27.1%; Score 13; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVEY 6  
 : ||  
 Db 3 VVEY 6

Search completed: April 22, 2003, 13:31:00  
 Job time : 24 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:27:31, Search time 28 seconds  
(without alignments)

66.229 Million cell updates/sec

Title: US-09-674-716B-7

Perfect score: 48

Sequence: 1 QQLVEYPFT 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 686

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL\_21.\*

1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	35.4	8	2	Q49534 mycoplasma
2	17	35.4	9	8	Q8WFT4 diadema ant
3	16	33.3	8	8	Q9MSX1 jurinea hum
4	16	33.3	9	6	Q9TRU7 bos taurus
5	15	31.2	8	4	Q15898 homo sapien
6	15	31.2	9	4	Q14277 homo sapien
7	15	31.2	9	4	Q96T78 homo sapien
8	15	31.2	9	12	Q65711 berne virus
9	14	29.2	8	2	Q934S4 thlobacillu
10	14	29.2	8	3	P82858 puccinia re
11	14	29.2	8	7	Q95213 oryctolagus
12	14	29.2	8	15	Q98YK9 human immun
13	14	29.2	9	4	Q9UC36 homo sapien
14	14	29.2	9	8	Q9MVU2 caloglossa
15	14	29.2	9	8	Q9TJ85 caloglossa
16	14	29.2	9	8	Q9TLD0 bostrychia

17	14	29.2	9	8	Q9MVU1 caloglossa
18	14	29.2	9	8	Q9TJ87 caloglossa
19	14	29.2	9	8	Q95BY3 caloglossa
20	14	29.2	9	8	Q8W7T9 bostrychia
21	14	29.2	9	8	Q9T388 caloglossa
22	14	29.2	9	8	Q9T387 bostrychia
23	14	29.2	9	8	Q78337 caloglossa
24	14	29.2	9	8	Q9T389 caloglossa
25	14	29.2	9	10	O81962 caloglossa
26	14	29.2	9	10	O81964 caloglossa
27	14	29.2	9	10	O81966 caloglossa
28	14	29.2	9	10	O81968 caloglossa
29	14	29.2	9	10	Q9S8J8 oryza sativ
30	14	29.2	9	10	O82778 caloglossa
31	14	29.2	9	11	Q9QWT0 mus musculu
32	13	27.1	7	8	O98866 spinacia ol
33	13	27.1	8	3	P87225 saccharomyc
34	13	27.1	8	4	Q9UJ50 homo sapien
35	13	27.1	8	4	Q9POK3 homo sapien
36	13	27.1	8	8	Q8WGC9 upogebia af
37	13	27.1	8	12	O83332 murine hepa
38	13	27.1	9	2	Q9R735 streptomyc
39	13	27.1	9	4	Q9UMF3
40	13	27.1	9	6	Q28121 bos taurus
41	13	27.1	9	13	Q9PRJ4 lepisosteus
42	13	27.1	9	13	P83060 bombina ori
43	13	27.1	9	13	P83059 bombina ori
44	13	27.1	9	13	P83058 bombina var
45	13	27.1	9	13	P83057 bombina var

#### ALIGNMENTS

##### RESULT 1

Q49534 ID Q49534 PRELIMINARY; PRT; 8 AA.  
AC Q49534;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE P120 (fragment).  
GN P120.  
OS Mycoplasma hominis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_taxid=2098;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V2785;  
RA Nyvold C., Birkelund S., Christiansen G.;  
RT "The Mycoplasma hominis P120 membrane protein gene contains a 659 base pair hypervariable domain."  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U22025; AAA67455.1; -.  
FT NON TER 1  
SQ SEQUENCE 8 AA; 869 MW; 914457605B02C05D CRC64;

Query Match 35.4%; Score 17; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 6.7e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVEYP 7  
Db 2 IVEIP 6

##### RESULT 2

Q8WFT4 ID Q8WFT4 PRELIMINARY; PRT; 9 AA.  
AC Q8WFT4;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN COII.  
 OS Diadema antillarum.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Diadematacea; Diadematoidea; Diademataidae;  
 OC Diadema.  
 OX NCBI\_TaxID=105358;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DCA3;  
 RX MEDLINE=21323357; PubMed=11430656;  
 RA Lessios H.A., Kessing B.D., Pearse J.S.;  
 RT "Population structure and speciation in tropical seas: global  
 RT phylogeography of the sea urchin *Diadema*.";  
 RL Evolution 55:955-975(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DCA3;  
 RX MEDLINE=21561594; PubMed=11703875;  
 RA Lessios H.A., Garrido M.J., Kessing B.D.;  
 RT "Demographic history of *Diadema antillarum*, a keystone herbivore on  
 RT Caribbean reefs.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).  
 DR EMBL; AY012858; AAL33832.2; -.  
 FT Mitochondrion.  
 KW NON\_TER 1  
 SQ SEQUENCE 9 AA; 1115 MW; EDF8DB1B173B46CA CRC64;  
  
 Query Match 35.4%; Score 17; DB 8; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 6.7e+05;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
 QY 1 QQLVEYP 7  
 DB 3 QYLEEQP 9  
  
 RESULT 3  
 ID QMSX1 PRELIMINARY; PRT; 8 AA.  
 AC QMSX1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE PABA (Fragment).  
 GN PSBA.  
 OS Jurinea humilis.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Asterales; Asteraceae; Cardueae; Jurinea.  
 OX NCBI\_TaxID=41594;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA O'Hanlon P.C., Briesse D.T.;  
 RT "Evidence for hierarchical and non-hierarchical evolution in the  
 RT Carduinae thistles.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF129846; AAF78138.1; -.  
 FT Chloroplast.  
 KW NON\_TER 1  
 SQ SEQUENCE 8 AA; 788 MW; 457451B5A76DDB10 CRC64;  
  
 Query Match 33.3%; Score 16; DB 8; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 4 VEYPT 9  
 DB 1 IEAPST 6

RESULT 4  
 ID Q9TRU7 PRELIMINARY; PRT; 9 AA.  
 AC Q9TRU7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE GAP-3, GTPase-activating protein (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92112868; PubMed=1309786;  
 RA Nice E.C., Fabri L., Hammacher A., Holden J., Simpson R.J.,  
 RA Burgess A.W.;  
 RT "The purification of a Rap1 GTPase-activating protein from bovine  
 RT brain cytosol.";  
 RL J. Biol. Chem. 267:1546-1553(1992).  
 FT NON\_TER 1  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1063 MW; 89EDA77B47604B5A CRC64;  
  
 Query Match 33.3%; Score 16; DB 6; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 4 VEYP 7  
 DB 4 IPYP 7  
  
 RESULT 5  
 ID Q15898 PRELIMINARY; PRT; 8 AA.  
 AC Q15898;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE (clone XP6A11B) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,  
 RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,  
 RA Caskey C.T.H.;  
 RT "Isolation of chromosome-specific genes by reciprocal probing of  
 RT arrayed cDNAs and cosmid libraries.";  
 RL Hum. Mol. Genet. 0:0-0(1995).  
 DR EMBL; L32078; AAA73888.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 938 MW; 34A415B0477B45BB CRC64;  
  
 Query Match 31.2%; Score 15; DB 4; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 6 YPFT 9  
 DB 3 YPIS 6  
  
 RESULT 6  
 ID Q14277 PRELIMINARY; PRT; 9 AA.

```
AC Q14277;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Homo sapiens short form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94071887; PubMed=7902707;
RA Ceccherini I., Boccardi R., Luo Y., Pasini B., Hofstra R.,
RA Takahashi M., Romeo G.;
RT "Exon structure and flanking intronic sequences of the human RET
RT proto-oncogene.";
RL Biochem. Biophys. Res. Commun. 196:1288-1295(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94366753; PubMed=8084609;
RA Ceccherini I., Hofstra R., Yin L., Stulp R., Barone V., Stelwagen T.,
RA Bocciaardi R., Nijveen H., Bolino A., Seri M., Ronchetto P., Pasini B.,
RA Bozzano M., Buys C., Romeo G.;
RT "DNA polymorphisms and conditions for SSCP analysis of the 20 exons of
RT the ret proto-oncogene.";
RL Oncogene 9:3025-3029(1994).
DR EMBL; U11532; AAC50102.1; -.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1134 MW; 034E11A9CDD1F5B0 CRC64;

Query Match 31.2%; Score 15; DB 4; Length 9;
Best Local Similarity 33.3%; Pred. No. 6.7e+05;
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 VEYFPT 9
DB 2 ISHAFT 7

RESULT 7
ID Q96T78 PRELIMINARY; PRT; 9 AA.
AC Q96T78;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Contactin-associated protein 2 (Fragment).
GN CNTNAP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21250995; PubMed=11352571;
RA Nakabayashi K., Scherer S.W.;
RT "The human contactin-associated protein 2 (CNTNAP2) spans over 2 Mb of
RT DNA at chromosome 7q35.";
RL Genomics 73:108-112(2001).
DR EMBL; AF318295; AAK49906.1; -.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 1080 MW; 9139A2D5A77B51EA CRC64;

Query Match 31.2%; Score 15; DB 4; Length 9;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 LVEYF 7
DB 1 LDHYP 5
```

```
RESULT 8
Q65711 PRELIMINARY; PRT; 9 AA.
AC Q65711;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF3 (Fragment).
OS Berne virus (BEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Torovirus.
OX NCBI_TaxID=111156;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9080137; PubMed=2293666;
RA Snijder E.J., Horzinek M.C., Spaan W.J.M.;
RT "A 3'-coterminally nested set of independently transcribed mRNAs is
RT generated during Berne virus replication.";
RL J. Virol. 64:331-338(1990).
DR EMBL; M33502; AAA42817.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1234 MW; D8EE736B5451AB19 CRC64;

Query Match 31.2%; Score 15; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 6.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YPF 8
DB 7 WPF 9

RESULT 9
Q934S4 PRELIMINARY; PRT; 8 AA.
AC Q934S4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MerD protein (Fragment).
GN MERD.
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=G66; TRANSPOSON=TN5037;
RA Kalyaeva E.S., Kholodii G.Y., Bass I.A., Gorlenko A.M., Yurieva O.V.,
RA Nikiforov V.G.;
RT "Tn5037, a Tn21-like mercury resistance transposon from Thiobacillus
RT ferrooxidans.";
RL Russ. J. Genet. 37:972-975(2001).
DR EMBL; AJ251743; CAC69252.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 937 MW; ED15A2D77B5DD446 CRC64;

Query Match 29.2%; Score 14; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YP 7
DB 4 YP 5

RESULT 10
P82858 PRELIMINARY; PRT; 8 AA.
AC P82858;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Uricase (EC 1.7.3.3) (Urate oxidase) (Fragment).
```

```

OS Puccinia recondita f. sp. triseti.
OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Urediniomycetidae;
OC Uredinales; Pucciniaceae; Puccinia.
OX NCBI_TaxID=142679;
RN [1]
RP SEQUENCE, FUNCTION, AND DEVELOPMENTAL STAGE.
RC TISSUE=SPORE;
RA Aguilar M., Montalbini P., Pineda M.;
RL Submitted (NOV-2000) to the SWISS-PROT data bank.
CC -1- FUNCTION: INVOLVED IN HOST-PARASITE RELATIONSHIP BETWEEN PLANTS
CC AND FUNGI.
CC -1- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE +
CC H(2)O(2) (5-HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: GERMINATION.
CC -1- SIMILARITY: BELONGS TO THE URICASE FAMILY.
DR InterPro: IPR002042; Uricase.
DR PROSITE: PS00366; URICASE; PARTIAL.
KW Oxidoreductase; Purine metabolism; Peroxisome.
FT NON_TER 8
SQ SEQUENCE 8 AA; 777 MW; 98CLADD735B9D76D CRC64;

Query Match 29.2%; Score 14; DB 3; Length 8;
Best Local Similarity 66.7%; Pred. No. 6.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 PFT 9
DB 2 PFS 4

RESULT 11
Q95213
ID Q95213 PRELIMINARY; PRT; 8 AA.
AC Q95213;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Germline DH (DH) gene (Fragment).
GN DH.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Lagonomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P-I/RGM;
RA Mage R.G., Chen H.-T., Alexander C.B., Chen F.F.;
RT "Rabbit DQ52 and DH Gene Rearrangements in Early B-cell Development.";
RL Mol. Immunol. 0:0-0(1996).
DR EMBL: U62585; AAB18735.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 845 MW; 5CA861B5A858677B CRC64;

Query Match 29.2%; Score 14; DB 7; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YP 7
DB 1 YP 2

RESULT 12
Q98YK9
ID Q98YK9 PRELIMINARY; PRT; 8 AA.
AC Q98YK9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Truncated pol protein (Fragment).
GN POL.

```

```

OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=991614;
RA Schmidt B., Walter H., Moschik G., Paatz C., Werwein M., Schwingel E.,
RA Korn K.;
RT "Recovery of HIV-1 pol gene sequences by direct sequencing of
RT amplification products derived from plasma samples.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF347458; AAK32535.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 845 MW; 72CDDDB1DD736CAB8 CRC64;

Query Match 29.2%; Score 14; DB 15; Length 8;
Best Local Similarity 75.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 OLVE 5
DB 3 QLAE 6

RESULT 13
Q9UC36
ID Q9UC36 PRELIMINARY; PRT; 9 AA.
AC Q9UC36;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 28 kDa heat shock protein homolog fragment 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RA MEDLINE=92218434; PubMed=1560006;
RA Kato K., Shinohara H., Goto S., Inaguma Y., Morishita R., Asano T.;
RT "Copurification of small heat shock protein with alpha B crystallin
RT from human skeletal muscle.";
RL J. Biol. Chem. 267:7718-7725(1992).
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 1220 MW; 26933415B1F77B43 CRC64;

Query Match 29.2%; Score 14; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YP 7
DB 3 YP 4

RESULT 14
Q9MVU2
ID Q9MVU2 PRELIMINARY; PRT; 9 AA.
AC Q9MVU2; Q9MVU3; Q9MVU4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit
DE (Fragment).
GN RBCL.
OS Caloglossa intermedia.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Delesseriaceae;
OC Caloglossa.
OX NCBI_TaxID=100879;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=SOUTH CAROLINA USA, NEW JERSEY USA, AND GEORGIA USA;
RA Kamiya M., West J.A., Zuccarello G.C., Kawai H.;
RT "Caloglossa intermedia sp. nov. (Delesseriaceae, Rhodophyta),
RT morphologically intermediate between C. lepieurii and C. monosticha,
RT from the western Atlantic coast.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030265; BAA90253.1; -
DR EMBL; AB030265; BAA90251.1; -
DR EMBL; AB030264; BAA90249.1; -
KW Chloroplast.
FT NON_TER 1
SQ SEQUENCE 9 AA; 977 MW; CAIA4DC1B771AB02 CRC64;

Query Match      29.2%; Score 14; DB 8; Length 9;
Best Local Similarity 75.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 VEYP 7
DB 2 VETP 5

RESULT 15
Q9TJ85 PRELIMINARY; PRT; 9 AA.
AC Q9TJ85;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit
DE (Fragment).
GN RBCL.
OS Caloglossa stipitata.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Delesseriaceae;
OC Caloglossa.
OX NCBI_TaxID=88391;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Kamiya M., Tanaka J., King R.J., West J.A., Zuccarello G.C., Kawai H.;
RT "Reproductive and genetic distinction between broad and narrow
RT entities of Caloglossa continua (Delesseriaceae, Rhodophyta).";
RL Phycologia 38:356-367(1999).
DR EMBL; AB023384; BAA88918.1; -
KW Chloroplast.
FT NON_TER 1
SQ SEQUENCE 9 AA; 977 MW; CAIA4DC1B771AB02 CRC64;

Query Match      29.2%; Score 14; DB 8; Length 9;
Best Local Similarity 75.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 VEYP 7
DB 2 VETP 5

```

Search completed: April 22, 2003, 13:31:37  
Job time : 30 secs





GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:24:36 ; Search time 73 Seconds  
(without alignments)  
16.428 Million cell updates/sec

Title: US-09-674-716B-7  
Perfect score: 48  
Sequence: 1 QQLVEYPT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 130868

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq 101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	9	21 AAY32256	Light chain CDR L3
2	44	91.7	9	19 AAW39823	Light chain CDR3 o
3	36	75.0	9	19 AAW39817	Light chain CDR3 o
4	34	70.8	9	21 AAY92171	Murine 13H10 light
5	33	68.8	9	14 AAR30450	C242:11 MAB kappa
6	31	64.6	9	21 AAY97234	Complementary dete
7	31	64.6	9	22 AAE13142	Humanised antibody
8	31	64.6	9	22 AAB82708	VEGF antagonist an
9	31	64.6	9	22 AAB82896	Anti-human CD154 a
10	31	64.6	9	22 AAG63993	Complementarity de

11	31	64.6	9	23 AAU74411	Light chain comple
12	30	62.5	9	19 AAW59175	Mouse PAF receptor
13	30	62.5	9	19 AAW39820	Light chain CDR3 o
14	30	62.5	9	19 AAW39826	Light chain CDR3 o
15	28	58.3	9	18 AAW24776	Human immunoglobul
16	28	58.3	9	19 AAW73171	CDR3 of light chal
17	28	58.3	9	19 AAW80146	Light chain CDR3 p
18	28	58.3	9	19 AAW62016	Light chain variab
19	28	58.3	9	20 AAY40428	Amino acid sequenc
20	28	58.3	9	20 AAY28392	Peptide fragment f
21	28	58.3	9	21 AAY82342	Humanised anti-CD1
22	28	58.3	9	22 AAB83165	Mouse ganglioside
23	27	56.2	8	14 AAR35879	Hepatitis C virus
24	27	56.2	8	14 AAR35880	Hepatitis C virus
25	27	56.2	8	14 AAR35878	Hepatitis C virus
26	27	56.2	9	18 AAW23435	CDR-3 of rW12 ligh
27	27	56.2	9	19 AAW39877	Light chain CDR3 o
28	27	56.2	9	20 AAW89158	Humanised antibody
29	27	56.2	9	21 AAY95235	Humanised antibody
30	27	56.2	9	22 AAB60407	Erbb2-reactive VL
31	27	56.2	9	22 AAB61592	Humanised Fab vers
32	26	54.2	7	12 AAR13525	Analgesic heptapep
33	26	54.2	7	16 AAR82702	Blood triglyceride
34	26	54.2	7	17 AAR97351	Peptide fragment o
35	26	54.2	7	18 AAW30321	Haemorphin peptide
36	26	54.2	7	18 AAW01800	Heptaepetide which
37	26	54.2	7	20 AAY42565	Oligopeptide compo
38	26	54.2	7	22 AAU79894	Anti-stress agent
39	26	54.2	7	23 AAU98728	Peptide inhibitor
40	26	54.2	8	17 AAR97350	Peptide fragment o
41	26	54.2	8	18 AAW30320	Haemorphin peptide
42	26	54.2	9	17 AAR97349	Peptide fragment o
43	26	54.2	9	18 AAW30319	Haemorphin peptide
44	26	54.2	9	19 AAW71857	Human anti-tissue
45	26	54.2	9	20 AAY06698	Ab2 variable light

## ALIGNMENTS

RESULT 1  
AAY32256  
ID AAY32256 standard; Peptide; 9 AA.  
XX AC AAY32256;  
XX DT 15-FEB-2000 (first entry)  
XX DE Light chain CDR L3 of mouse anti-CD23 MAB C11.  
XX KW CD23; FCERII; IGE receptor; monoclonal antibody; C11; mouse;  
KW monoclonal antibody; chimeric antibody; humanised antibody;  
KW complementarity determining region; CDR; autoimmune disease;  
KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
KW urticaria; nephrotic syndrome; glomerulonephritis;  
KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
KW graft-versus-host disease; COPD; bronchitis; diabetes;  
KW B-cell malignancy; therapy.  
XX OS Mus musculus.  
XX PN WO958679-A1.  
XX PD 18-NOV-1999.  
XX PF 07-MAY-1999; 99WO-GB01434.  
XX PR 09-MAY-1998; 98GB-0009839.  
XX PA (GLAX ) GLAXO GROUP LTD.  
XX XX

PI Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 XX WPI; 2000-053101/04.  
 DR N-PSDB; AAZ34741.  
 XX  
 PT Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 XX  
 PS Claim 1; Page 40; 81pp; English.  
 XX  
 CC This sequence represents complementarity determining region 3  
 CC (CDR L3) of the light chain of murine anti-CD23 (PCERII) monoclonal  
 CC antibody C11 (see also AAY32262). The invention provides altered  
 CC antibodies, such as chimeric or humanised antibodies, which comprise  
 CC sufficient of the amino acid sequences of C11 light and heavy chain  
 CC CDRs (see AAY32254-59) to render them capable of binding to the CD23  
 CC type II molecule expressed on haematopoietic cells. The antibodies  
 CC are used to block soluble CD23 formation for treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
 |||||  
 Db 1 QQLVEYPFT 9

RESULT 2  
 AAW39823  
 ID AAW39823 standard; peptide; 9 AA.  
 XX  
 AC AAW39823;  
 XX  
 DT 16-JUN-1998 (first entry)  
 XX  
 DE Light chain CDR3 of catalytic antibody 2A10.  
 XX  
 KW Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9749800-A1.  
 XX  
 PD 31-DEC-1997.  
 XX  
 PF 25-JUN-1997; 97WO-US10965.  
 XX  
 PR 25-JUN-1996; 96US-0672345.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Landry DW;  
 XX  
 DR WPI; 1998-077166/07.  
 XX  
 PT New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 XX  
 PT

PT\* in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX  
 PS Claim 15; Page 82; 147pp; English.  
 XX  
 CC AAW39821-23 represent the sequences of the light chain complementarity  
 CC determining regions (CDRs) of the catalytic antibody 2A10, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 2A10 antibody was  
 CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
 CC monoester transition state analogue. Antibody 2A10 has a per minute Kcat  
 CC of 0.011. The antibodies reduce the concentration of cocaine in a  
 CC subject, and are used particularly for the treatment of an overdose. They  
 CC are also used for treating addiction (by reducing the in vivo  
 CC concentration that can be achieved).  
 XX  
 SQ Sequence 9 AA;

Query Match 91.7%; Score 44; DB 19; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 7.8e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
 |||||  
 Db 1 QQFVEYPFT 9

RESULT 3  
 AAW39817  
 ID AAW39817 standard; peptide; 9 AA.  
 XX  
 AC AAW39817;  
 XX  
 DT 16-JUN-1998 (first entry)  
 XX  
 DE Light chain CDR3 of catalytic antibody 3B9.  
 XX  
 KW Variable domain; lambda light chain; catalytic antibody; degradation;  
 KW cocaine; cocaine transition state analogue; TSA; benzoic acid;  
 KW phenyl cocaine; immunogenic conjugate; reduction; cocaine; treatment;  
 KW overdose; addiction.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9749800-A1.  
 XX  
 PD 31-DEC-1997.  
 XX  
 PF 25-JUN-1997; 97WO-US10965.  
 XX  
 PR 25-JUN-1996; 96US-0672345.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Landry DW;  
 XX  
 DR WPI; 1998-077166/07.  
 XX  
 PT New catalytic antibodies able to decompose cocaine, single-chain  
 PT analogues - used to treat cocaine overdose and addiction, required  
 PT in far smaller doses than antibodies that antagonise cocaine by  
 PT simply binding  
 XX  
 PS Claim 11; Page 81; 147pp; English.  
 XX  
 CC AAW39815-17 represent the sequences of the light chain complementarity  
 CC determining regions (CDRs) of the catalytic antibody 3B9, which is able  
 CC to degrade cocaine. A series of cocaine transition state analogues  
 CC (TSAs) were prepared and used to immunise mice for production of  
 CC hybridomas. Catalytic antibodies were identified by their capacity to  
 CC release 3H-benzoic acid from 3H-phenyl cocaine. The 3B9 antibody was

CC identified using TSA1, which is an immunogenic conjugate of a phosphate  
 CC monoeaster transition state analogue. Antibody 3B9 has a per minute Kcat  
 CC of 0.11. The antibodies reduce the concentration of cocaine in a subject,  
 CC and are used particularly for the treatment of an overdose. They are also  
 CC used for treating addiction (by reducing the in vivo concentration that  
 CC can be achieved).

XX Sequence 9 AA;

Query Match 75.0%; Score 36; DB 19; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QLVVEYPPT 9

Db 1 QHFVDYPT 9

RESULT 4

AAV92171  
 ID AAY92171 standard; Peptide; 9 AA.

XX  
 AC AAY92171;

DT 01-AUG-2000 (first entry)

DE Murine 13H10 light chain variable region CDR 3.

XX Light chain; variable region; complementarity determining region; CDR 3;  
 KW anti-Tie2 kinase receptor; monoclonal antibody; 1598; angiogenetic;  
 KW vascular-general; proliferative; antischismic; cerebroprotective;  
 KW cardiant; agonist; antibody inhibition.

XX Mus musculus.

OS WO200018804-A1.

PN 06-APR-2000.

XX 28-SEP-1999; 99WO-US22428.

XX 28-SEP-1998; 98US-0102098.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX Holmes SD, Erickson-miller CL, Winkler JD;

PI WPI; 2000-293114/25.

XX Tie2 receptor agonist antibodies useful for promoting angiogenesis in  
 PT patients suffering from strokes and myocardial infarctions

PS Claim 20; Page 46; 50pp; English.

XX AAY92169-71 are light chain CDR (complementarity determining regions)  
 CC from a novel murine anti-Tie2 kinase receptor agonist monoclonal  
 CC antibody 13H10. Tie2 is a single-transmembrane, tyrosine kinase receptor  
 CC ('Tie' stands for tyrosine kinase receptor with immunoglobulin and  
 CC endothelial growth factor (EGF) homology domain(s)). Anti-Tie2  
 CC antibodies may be administered to enhance angiogenesis in mammals  
 CC suffering from ischemic disease, myocardial infarction or cerebral stroke  
 CC or other vascular diseases such as diabetes. It may also be used to  
 CC enhance endothelial cell survival and to promote hematopoietic or  
 CC megakaryocyte cell proliferation (claimed).

XX Sequence 9 AA;

Query Match 70.8%; Score 34; DB 21; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QLVVEYPPT 9

Db 2 QLVVEYPPT 9

RESULT 5

AAAR30450  
 ID AAR30450 standard; peptide; 9 AA.

XX  
 AC AAR30450;

DT 06-MAY-1993 (first entry)

DE C242:11 Mab kappa chain CDR3.

XX Kappa; chain; heavy; complementarity determining region; CDR; Mab;  
 KW monoclonal antibody; C242:11; murine; IgG; hybridoma; cell line;  
 KW spleen; human; colonic; adenocarcinoma; myeloma; Sp2/0; antigen;  
 KW endocytosis.

OS Synthetic.

PN EF521842-A.

XX 07-JAN-1993.

PF 03-JUL-1992; 92EP-0850166.

PR 03-JUL-1991; 91SE-0002074.

XX (KABI ) KABI PHARMACIA AB.

PI Holmgren J, Lind P, Lindholm L;

DR WPI; 1993-002345/01.

XX Monoclonal antibody reacting with CA-242 antigen - obt'd. by  
 PT culturing hybridoma cell line C242:11 or mutants, useful for  
 PT diagnosis and therapy of pancreatic or colorectal cancers

PS Claim 5; Page 11; 15pp; English.

XX The sequences given in AAR30448-50 represent the kappa chain and  
 CC AAR30451-53 the heavy chain complementarity determining regions (CDR)  
 CC of a monoclonal antibody (Mab), C242:11. C242:11 is a monoclonal  
 CC murine Ab of IgG class produced when culturing in an appropriate  
 CC medium a hybridoma cell line obtained by fusing spleen cells from a  
 CC mouse, which has been immunised with a human colonic adenocarcinoma  
 CC cell line, with the murine myeloma cell line Sp2/0. C242:11 when  
 CC bound to a cell surface antigen is capable of being endocytosed or  
 CC internalised into cells.

XX Sequence 9 AA;

Query Match 68.8%; Score 33; DB 14; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QLVVEYPPT 9

Db 2 QLVVEYPPT 9

RESULT 6

AAAY97234  
 ID AAY97234 standard; Protein; 9 AA.

XX  
 AC AAY97234;

DT 19-DEC-2000 (first entry)

DE Complementary determining region (CDRL3) of anti-SI (KDR) antibody.

XX Immunoglobulin; antibody; complementary determining region; CDR;

KW VEGF; vascular endothelial growth factor; KDR;  
 KW kinase insert domain containing receptor; multivalent; monovalent;  
 KW humanised antibody; chimeric antibody; tumour; diabody; triabody;  
 KW glioblastoma multiforme; hemangioblastoma; AIDS;  
 KW central nervous system neoplasms; AIDS associated Kaposi's sarcoma;  
 KW acquired immune deficiency syndrome; AIDS; human.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200044777-A1.  
 XX  
 PD 03-AUG-2000.  
 XX  
 PF 28-JAN-2000; 2000WO-US02180.  
 XX  
 PR 29-JAN-1999; 99US-0117726.  
 PR 29-JAN-1999; 99US-0240736.  
 XX  
 PA (IMCL-) IMCLONE SYSTEMS INC.  
 PI Zhu Z, Witte L;  
 XX  
 DR WPI; 2000-505366/45.  
 DR N-PSDB; AAA53766.  
 XX  
 PT Novel immunoglobulin molecules binding kinase insert domain-containing  
 PT receptor with the same affinity as vascular endothelial growth factor,  
 PT used to reduce tumour growth  
 XX  
 PS Claim 3; Page 50; 55pp; English.  
 CC New immunoglobulin molecules are described that bind kinase insert  
 CC domain-containing receptor (KDR) with a comparable affinity to human  
 CC vascular endothelial growth factor (VEGF). The antibodies neutralise  
 CC KDR activation. The immunoglobulin may be a multivalent single  
 CC chain antibody, a monovalent single chain antibody, a diabody, a  
 CC triabody, a humanised antibody or a chimerised antibody.  
 CC The immunoglobulin molecules bind specifically to an  
 CC extracellular domain of the KDR receptor with the same affinity as  
 CC VEGF. Overexpression of the KDR receptor with the same affinity as  
 CC human tumour cell lines including glioblastoma multiforme,  
 CC hemangioblastoma, central nervous system neoplasms and AIDS  
 CC associated Kaposi's sarcoma. The antibodies therefore have  
 CC applications in treating these conditions. This sequence encodes a  
 CC preferred heavy chain complementary determining region of the  
 CC immunoglobulins of the invention.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 64.6%; Score 31; DB 21; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 QQLVEYPPT 9  
 |||||  
 Db 1 QQRSSYPPT 9  
 |||||  
 RESULT 7  
 AAEL13142  
 ID AAEL13142 standard; peptide; 9 AA.  
 XX  
 AC AAEL13142;  
 XX  
 DT 28-JAN-2002 (first entry)  
 XX  
 DE Humanised antibody murine light chain hypervariable region (VL) CDR3.  
 XX  
 KW Vascular endothelial growth factor receptor; VEGFR; antagonist; tumour;  
 KW cytosolic; light chain hypervariable region; VL; myelocytic leukaemia;  
 KW lymphocytic; erythrocytic; monocytic; multiple myeloma; lymphoid cell;  
 KW Hodgkin's disease; complementarity determining region-3; CDR-3; mouse.

XX Mus sp.  
 OS  
 XX WO200174296-A2.  
 PN  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US10504.  
 XX  
 PR 31-MAR-2000; 2000US-0540770.  
 XX  
 PA (IMCL-) IMCLONE SYSTEMS INC.  
 PA (CORR ) CORNELL RES FOUND INC.  
 XX  
 PI Witte L, Rafii S;  
 XX  
 DR WPI; 2001-662942/76.  
 DR N-PSDB; AAD21668.  
 XX  
 PT Inhibiting growth of non-solid tumor cells useful to treat bone marrow  
 PT tumors such as leukemias or multiple myeloma comprises treatment with  
 PT an antagonist of a vascular endothelial growth factor receptor -  
 XX  
 PS Claim 8; Page 15; 68pp; English.  
 CC The invention relates to a method for inhibiting the growth of non-solid  
 CC tumour cells that are stimulated by a ligand of vascular endothelial  
 CC growth factor receptor (VEGFR) in mammals particularly humans. The method  
 CC involves treating the mammals with humanised VEGFR monoclonal antibodies  
 CC (antagonists). Humanised monoclonal antibody comprises humanised mouse  
 CC variable region joined to human constant region, where the humanised  
 CC mouse variable region contains mouse complementarity determining region  
 CC (CDR) grafted into human variable region. The method is useful for  
 CC treating leukaemias such as acute or chronic myelocytic leukaemia, acute  
 CC or chronic lymphocytic leukaemia, erythrocytic or monocytic leukaemia,  
 CC multiple myelomas and lymphoid cells, particularly those related to  
 CC non-Hodgkin's and Hodgkin's disease. The present sequence is humanised  
 CC antibody murine light chain hypervariable region (VL) CDR-3 used in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 64.6%; Score 31; DB 22; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 QQLVEYPPT 9  
 |||||  
 Db 1 QQRSSYPPT 9  
 |||||  
 RESULT 8  
 AAB82708  
 ID AAB82708 standard; Peptide; 9 AA.  
 XX  
 AC AAB82708;  
 XX  
 DT 15-OCT-2001 (first entry)  
 XX  
 DE VEGF antagonist antibody IMC-1C11 VL CDR-3.  
 XX  
 KW IMC-1C11; chimeric antibody; mouse; human; antagonist; VEGF;  
 KW vascular endothelial growth factor; angiogenesis; antiangiogenic;  
 KW antitumour; neoplasm; collagen-vascular disease; autoimmune disease;  
 KW tumour; breast carcinoma; lung carcinoma; prostate carcinoma;  
 KW colon carcinoma; ovarian carcinoma; neuroblastoma;  
 KW glioblastoma multiforme; melanoma; therapy; light chain; CDR;  
 KW complementarity determining region.  
 XX  
 OS Chimeric - Mus sp.  
 OS Chimeric - Homo sapiens.  
 XX  
 PN WO200154723-A1.

XX PD 02-AUG-2001.  
 XX PF 29-JAN-2001; 2001WO-US02839.  
 XX PR 28-JAN-2000; 2000US-0178791.  
 XX PR 31-MAR-2000; 2000US-0539692.  
 XX (SUNN-) SUNNYBROOK HEALTH SCI CENT.  
 PA (IMCL-) IMCLONE SYSTEMS INC.  
 XX Kerbel R;  
 XX WPI; 2001-514531/56.  
 XX Treating or controlling an angiogenic dependent condition (e.g. a  
 PT neoplasm, collagen-vascular or autoimmune disease) in mammal by  
 PT administering a combination of an antiangiogenic molecule and a  
 PT chemotherapeutic agent  
 XX Disclosure; Page 37; 42pp; English.  
 XX The present sequence is that of complementarity determining region  
 CC 3 of the light chain variable region (see also AAB82702) of  
 CC IMC-1C11, a mouse-human chimeric antibody that has vascular  
 CC endothelial growth factor (VEGF) antagonist activity. The antibody,  
 CC or a fragment of it, can be used as an anti-angiogenic molecule,  
 CC together with a chemotherapeutic agent, for the treatment of an  
 CC angiogenic dependent condition in a mammal, especially a human.  
 CC The invention relates generally to a method of treating or  
 CC controlling an angiogenic dependent condition by administering an  
 CC anti-angiogenic molecule and a chemotherapeutic agent, to produce a  
 CC regression or arrest of the condition while minimising or  
 CC preventing significant toxicity of the chemotherapeutic agent.  
 CC The anti-angiogenic molecule inhibits or blocks the action of a  
 CC vascular endothelium survival factor such as VEGF or its receptor,  
 CC and is especially IMC-1C11. Conditions that can be treated include  
 CC a neoplasm, a collagen-vascular disease or an autoimmune disease,  
 CC especially a solid tumour, including breast carcinoma, lung  
 CC carcinoma, prostate carcinoma, colon carcinoma, prostate carcinoma,  
 CC ovarian carcinoma, neuroblastoma, central nervous system tumour,  
 CC neuroblastoma, glioblastoma multiforme or melanoma (all claimed).  
 XX Sequence 9 AA;  
 SQ Query Match 64.6%; Score 31; DB 22; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 QQLVEYPFT 9  
 DB 1 QQRSSYPFT 9  
 RESULT 9  
 AAB82896  
 ID AAB82896 standard; Peptide; 9 AA.  
 XX AC AAB82896;  
 XX DT 26-NOV-2001 (first entry)  
 XX Anti-human CD154 antibody ABI793 light chain variable region CDR3.  
 XX CD154; gp39; CD40-L; antibody; ABI793; transplant rejection;  
 KW autoimmune disease; inflammation; atherosclerosis;  
 KW Alzheimer's disease; antiinflammatory; antiarteriosclerotic;  
 KW immunosuppressive; therapy; complementarity determining region;  
 KW CDR3.  
 XX Mus musculus.  
 OS WO200168860-A1.  
 XX PN

XX PD 20-SEP-2001.  
 XX PF 14-MAR-2001; 2001WO-EP02875.  
 XX PR 16-MAR-2000; 2000GB-0006398.  
 XX (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 XX Di Padova FE, Schuler W;  
 XX WPI; 2001-590062/66.  
 XX CD154 binding molecule, in particular antibody to human CD154 for use  
 PT in treatment, prevention of autoimmune, inflammatory diseases,  
 PT atherosclerosis, Alzheimer's disease and prevention of transplant  
 PT rejection  
 XX Claim 2; Page 23; 37pp; English.  
 XX The present sequence is that of complementarity determining region  
 CC 3 (CDR3) of the light chain variable region (VL, see also AAB82890)  
 CC of mouse anti-human CD154 monoclonal antibody ABI793. The invention  
 CC provides a CD154 binding molecule, in particular an antibody to  
 CC human CD154, in which the VH region has the CDR1, CDR2 and CDR3 of  
 CC ABI793 VH and the VL region has the CDR1, CDR2 and CDR3 of ABI793  
 CC VL. The CD154 binding molecule is especially a human antibody in  
 CC which the VH and VL domains are essentially those of ABI793, with  
 CC constant regions from human heavy and light chains. The antibodies  
 CC are used to inhibit an immune response mediated by CD154-positive  
 CC cell interactions with CD40-positive cells, in the treatment and/or  
 CC prevention of diseases, disorders or conditions where CD154  
 CC modulation and/or interference with or inhibition of the CD154:CD40  
 CC interactions is therapeutically beneficial, prevention of  
 CC macrophage-associated inflammatory processes and in the treatment of  
 CC diseases where suppression of antibody responses to antigens is  
 CC desirable (claimed). The CD154 binding molecules are useful for  
 CC prevention of cell, tissue or organ graft rejection, in the prevention  
 CC and treatment of autoimmune or inflammatory diseases, atherosclerosis  
 CC or Alzheimer's disease, and also for inhibiting B cell proliferation  
 CC and differentiation, T cell responses, induction or modulation of T  
 CC or B cell tolerance or inhibition of the growth of tumour cells  
 CC expressing CD154 antigen. Such conditions include autoimmune and  
 CC non-autoimmune disorders, in particular, Addison's disease, Celiac  
 CC sprue, glomerulonephritis, Grave's disease, Hashimoto's thyroiditis,  
 CC haemolytic disease of the newborn, keratitis, multiple sclerosis,  
 CC polymyositis, psoriasis, rheumatic fever, rheumatoid arthritis,  
 CC sarcoidosis, syphilis, tuberculosis, ulcerative colitis,  
 CC HIV infection, leukaemia or lymphoma.  
 XX Sequence 9 AA;  
 SQ Query Match 64.6%; Score 31; DB 22; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 QQLVEYPFT 9  
 DB 1 QQYNSYPFT 9  
 RESULT 10  
 AAG63993  
 ID AAG63993 standard; peptide; 9 AA.  
 XX AC AAG63993;  
 XX DT 26-NOV-2001 (first entry)  
 XX Complementarity determining region of light chain of antibody 2C4.  
 DE Monoclonal antibody 2C4; sialoadhesin factor-2; SAF-2; allergic rhinitis;  
 KW

KW allergy; asthma; anemia; eczema; lymphoma; systemic mastocytosis;  
 KW leukemia; eosinophil.  
 OS Mus sp.  
 XX WO200166126-A1.  
 XX 13-SEP-2001.  
 XX 05-MAR-2001; 2001WO-US07193.  
 XX 07-MAR-2000; 2000US-0187595.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Abrahamson JA, Bochner B, Erickson-Miller CL, Kikly KK;  
 PI Schleimer R;  
 DR WPI; 2001-570749/64.  
 XX  
 PT Novel monoclonal antibody specific for human sialoadhesin factor-2 for  
 PT diagnosis, prevention, treatment of allergy, asthma, eczema or diseases  
 PT such as lymphoma, leukemia or systemic mastocytosis, in a mammal -  
 XX  
 PS Claim 10; Page 34; 35pp; English.  
 XX  
 CC AAG33901-93 represent the complementarity determining regions (CDRs)  
 CC of the light chain variable region of murine monoclonal antibody 2C4.  
 CC This antibody binds to human sialoadhesin factor-2 (SAF-2). The  
 CC antibody is useful for treating or preventing allergic rhinitis,  
 CC allergies, asthma, anemia, eczema or diseases such as lymphoma,  
 CC leukemia or systemic mastocytosis in a mammal. It is also useful for  
 CC detecting the presence of a cell, especially eosinophil in a sample,  
 CC by detecting binding of the antibody to SAF-2. The antibody can be  
 CC coupled to toxins, antiproliferative drugs or radionuclides to  
 CC kill cells in areas of excessive SAF-2 expression.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 64.6%; Score 31; DB 22; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 QQLVEYPFT 9  
 |||  
 DB 1 QQRSSYPFT 9  
 |||  
 RESULT 11  
 AAU74411  
 ID AAU74411 standard; peptide; 9 AA.  
 AC AAU74411;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Light chain complementarity determining region L3 (CDRL3).  
 XX  
 KW Complementarity determining region; CDR; CDRL3; antigen; cytostatic;  
 KW angiogenesis inhibitor; vascular endothelial growth factor receptor;  
 KW VEGF; tumour; leukaemia; antibody; cell proliferation inhibitor;  
 KW antibody light chain variable domain.  
 XX  
 OS Mus sp.  
 XX  
 XX WO200190192-A2.  
 PN 29-NOV-2001.  
 PD  
 XX 24-MAY-2001; 2001WO-US16924.  
 PF  
 XX

PR 24-MAY-2000; 2000US-206749P.  
 XX (IMCL-) IMCLONE SYSTEMS INC.  
 XX  
 PI Zhu Z;  
 XX  
 DR WPI; 2002-106189/14.  
 DR N-PSDB; AAS20282.  
 XX  
 XX New bispecific immunoglobulin-like antigen-binding protein for reducing  
 PT tumour growth and for inhibiting angiogenesis, comprises a complex of  
 PT two polypeptides and two second polypeptides -  
 XX  
 PS Claim 55; Page 57; 64pp; English.  
 XX  
 CC The invention describes an antigen-binding protein (I) comprising a  
 CC complex of two polypeptides (P1) and two second polypeptides (P2) which  
 CC are stably associated in an immunoglobulin like complex. P1 has an  
 CC antigen-binding site located to the N terminus of immunoglobulin (Ig)  
 CC light chain constant domain (CL domain), and P2 has an antigen-binding  
 CC site located to the N terminus of the CH1 domain. (I) is useful for:  
 CC neutralising the activation of a vascular endothelial growth factor  
 CC (VEGF) receptor; reducing tumour growth; inhibiting angiogenesis;  
 CC reducing endothelial cell proliferation; inhibiting VEGF induced  
 CC migration of human leukaemia cells; blocking interaction of a protein and  
 CC its ligand; promoting interactions between immune cells and target cells;  
 CC and in vivo and in vitro for investigative, diagnostic or treatment  
 CC methods. The design of (I) provides for efficient production so that  
 CC substantially all of the antigen-binding proteins produced are assembled  
 CC in the desired configuration. (I) is bivalent and bispecific, homogeneous  
 CC and in tetrameric form. The heavy chain constant domains which constitute  
 CC the Fc region (e.g., CH2 and CH3 for an Igg molecule) of a natural  
 CC antibody and which provide other antibody functions can be present. There  
 CC is no requirement for processing in vitro to obtain the complete product.  
 CC This peptide sequence represents the light chain variable domain  
 CC complementarity determining region L3 (CDRL3) incorporated into an  
 CC antigen-binding protein described in the method of the invention.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 64.6%; Score 31; DB 23; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 QQLVEYPFT 9  
 |||  
 DB 1 QQRSSYPFT 9  
 |||  
 RESULT 12  
 AAWS9175  
 ID AAWS9175 standard; Protein; 9 AA.  
 XX  
 AC AAWS9175;  
 XX  
 DT 13-AUG-1998 (first entry)  
 XX  
 DE Mouse PAF receptor antibody light chain hypervariable region CDR3.  
 XX  
 KW Immunoglobulin light chain; hypervariable region; PAF; CDR; antibody;  
 KW platelet activating factor receptor; murine; chimeric; anti-PAF;  
 KW complementarity determining region.  
 XX  
 OS Mus sp.  
 XX  
 XX JP10136981-A.  
 PN 26-MAY-1998.  
 PD  
 XX 12-NOV-1996; 96JP-0317047.  
 PF  
 XX 12-NOV-1996; 96JP-0317047.  
 PR  
 XX

(SUME ) SUMITOMO ELECTRIC IND CO.  
WPI; 1998-355043/31.

Polypeptide comprising variable region of antibody recognising PAF receptor - useful for the preparation of chimeric antibody

Claim 2; Page 10; 14pp; Japanese.

This sequence represents the complementarity determining region (CDR) 3 of the light chain hypervariable region of an antibody recognising the platelet activating factor (PAF) receptor. The variable region of mouse anti-PAF receptor monoclonal antibody is useful for the preparation of chimeric antibodies.

SQ Sequence 9 AA;

Query Match  
Best Local Similarity 62.5%; Score 30; DB 19; Length 9;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QQLVEYPPT 9  
| : ||||  
Db 1 QQYNTYPEFT 9

RESULT 13  
AAW39820  
ID AAW39820 standard; peptide; 9 AA.

AC AAW39820;

Dt 16-JUN-1998 (first entry)

Light chain CDR3 of catalytic antibody 6A12.

Variable domain; lambda light chain; catalytic antibody; degradation; cocaine; cocaine transition state analogue; TSA; benzoic acid; phenyl cocaine; immunogenic conjugate; reduction; cocaine treatment; overdose; addition.

Mus sp.

WO9749800-A1.

31-DEC-1997.

25-JUN-1997; 97WO-US10965.

25-JUN-1996; 96US-0672345.

(UYCO ) UNIV COLUMBIA NEW YORK.

Lantry DW;

New catalytic antibodies able to decompose cocaine, single-chain analogues - used to treat cocaine overdose and addiction, required in far smaller doses than antibodies that antagonise cocaine by simply binding

Claim 13; Page 82; 147pp; English.

AAW39818-20 represent the sequences of the light chain complementarity determining regions (CDRs) of the catalytic antibody 6A12, which is able to degrade cocaine. A series of cocaine transition state analogues (TSAs) were prepared and used to immunise mice for production of hybridomas. Catalytic antibodies were identified by their capacity to release 3H-benzoic acid from 3H-phenyl cocaine. The 6A12 antibody was identified using TSA1, which is an immunogenic conjugate of a phosphate monoester transition state analogue. Antibody 6A12 has a per minute Kcat of 0.072. The antibodies reduce the concentration of cocaine in a

CC subject, and are used particularly for the treatment of an overdose. They are also used for treating addiction (by reducing the in vivo CC concentration that can be achieved).

XX SQ Sequence 9 AA;

Query Match  
Best Local Similarity 62.5%; Score 30; DB 19; Length 9;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QQLVEYPPT 9  
| : ||||  
Db 1 QHFEDYPPT 9

RESULT 14  
AAW39826  
ID AAW39826 standard; peptide; 9 AA.

AC AAW39826;

Dt 16-JUN-1998 (first entry)

Light chain CDR3 of catalytic antibody 12H1.

Variable domain; lambda light chain; catalytic antibody; degradation; cocaine; cocaine transition state analogue; TSA; benzoic acid; phenyl cocaine; immunogenic conjugate; reduction; cocaine treatment; overdose; addition.

Mus sp.

WO9749800-A1.

31-DEC-1997.

25-JUN-1997; 97WO-US10965.

25-JUN-1996; 96US-0672345.

(UYCO ) UNIV COLUMBIA NEW YORK.

Lantry DW;

New catalytic antibodies able to decompose cocaine, single-chain analogues - used to treat cocaine overdose and addiction, required in far smaller doses than antibodies that antagonise cocaine by simply binding

Claim 17; Page 83; 147pp; English.

AAW39824-26 represent the sequences of the light chain complementarity determining regions (CDRs) of the catalytic antibody 12H1, which is able to degrade cocaine. A series of cocaine transition state analogues (TSAs) were prepared and used to immunise mice for production of hybridomas. Catalytic antibodies were identified by their capacity to release 3H-benzoic acid from 3H-phenyl cocaine. The 12H1 antibody was identified using TSA2, and has a per minute Kcat of 0.16. The antibodies reduce the concentration of cocaine in a subject, and are used particularly for the treatment of an overdose. They are also used for treating addiction (by reducing the in vivo concentration that can be achieved).

XX SQ Sequence 9 AA;

Query Match  
Best Local Similarity 62.5%; Score 30; DB 19; Length 9;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QQLVEYPPT 9  
| : ||||  
Db 1 QHFEDYPPT 9

CC subject, and are used particularly for the treatment of an overdose. They are also used for treating addiction (by reducing the in vivo CC concentration that can be achieved).

XX SQ Sequence 9 AA;

Query Match  
Best Local Similarity 62.5%; Score 30; DB 19; Length 9;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QQLVEYPPT 9  
| : ||||  
Db 1 QHFEDYPPT 9

```

Db      1 QHFEDYPFT 9

RESULT 15
AAW24776
ID AAW24776 standard; peptide; 9 AA.
XX
AC AAW24776;
XX
DT 03-DEC-1997 (first entry)
XX Human immunoglobulin light chain CDR3 region peptide.
DE
XX
XX Ig; affinity constant; human; antigen; hybridoma; B cell; transgene;
KW transgenic; mouse; CD4; antibody; autoimmune; inflammatory;
KW transplant rejection.
XX
XX Homo sapiens.
OS
XX
XX WO9713852-A1.
PN
XX
PD 17-APR-1997.
XX
PF 10-OCT-1996; 96WO-US16433.
XX
PR 10-OCT-1995; 95US-0544404.
XX
PA (GENP-) GENPHARM INT INC.
XX
PI Kay RM, Lonberg N;
XX
XX WPI; 1997-235888/21.
DR
XX
PT Novel anti-CD4 antibody produced by transgenic mice - used in the
PT treatment of auto-immune disease etc.
XX
PS Claim 60; Page 287; 396pp; English.
XX
CC A novel composition has been developed which comprises an immunoglobulin
CC (Ig) having an affinity constant (Ka) of at least 2 multiply
CC 1000000000 M-1 for binding to a predetermined human antigen. The
CC present sequence represents a human light chain CDR3 region peptide
CC which forms part of an immunoglobulin comprising a VKL15 segment,
CC a JK2 segment, and the present light chain CDR3 region. The
CC anti-CD4 antibodies may be used in therapeutic and diagnostic
CC applications, especially for the treatment of human diseases. These
CC antibodies reduce activity of CD4 cells and reduce undesirable
CC autoimmune reactions, inflammatory response and transplant rejection.
CC Transgenic animals are capable of producing heterologous antibodies
CC of multiple isotypes by undergoing isotype switching. These animals
CC produce a first Ig type that is necessary for antigen-stimulated B-cell
CC maturation and can switch to encode and produce one or more subsequent
CC heterologous isotypes.
XX
SQ Sequence 9 AA;

Query Match 58.3%; Score 28; DB 18; Length 9;
Best Local Similarity 55.6%; Pred. No. 7.8e+05;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9
|||
Db 1 QQYDSYPFT 9

Search completed: April 22, 2003, 13:30:30
Job time : 74 secs

```



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:31:41 ; Search time 41 Seconds  
(without alignments)  
17.589 Million cell updates/sec

Title: US-09-674-716B-7

Perfect score: 48

Sequence: 1 QQLVEYPT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 29135

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08 NEW PUB.pcp.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pcp.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06 NEW PUB.pcp.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07 NEW PUB.pcp.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09 NEW PUB.pcp.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10 NEW PUB.pcp.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60 NEW PUB.pcp.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	64.6	9	10	US-09-976-787-6
2	31	64.6	9	10	US-09-865-198-6
3	30	62.5	9	10	US-09-828-708-51
4	27	56.2	9	9	US-10-091-236-16
5	27	56.2	9	9	US-10-032-482-14
6	27	56.2	9	9	US-10-091-313-6
7	27	56.2	9	9	US-09-155-106-6
8	27	56.2	9	9	US-10-091-268-6
9	26	54.2	9	9	US-09-293-854-7
10	26	54.2	9	9	US-09-996-288-16
11	26	54.2	9	9	US-09-977-797A-62
12	26	54.2	9	9	US-09-977-797A-66
13	26	54.2	9	10	US-09-808-037-17
14	26	54.2	9	10	US-09-808-037-19
15	25	52.1	4	9	US-09-994-078-2
16	25	52.1	9	9	US-09-771-415-5
17	25	52.1	9	9	US-09-771-415-14
18	25	52.1	9	9	US-09-771-415-15
19	25	52.1	9	9	US-09-771-415-16

20	52.1	9	9	US-09-996-288-6	Sequence 6, Appli
21	52.1	9	9	US-09-996-288-61	Sequence 61, Appl
22	52.1	9	9	US-10-144-644-47	Sequence 47, Appl
23	52.1	9	9	US-10-144-644-48	Sequence 48, Appl
24	52.1	9	10	US-09-796-848A-7	Sequence 7, Appli
25	52.1	9	10	US-09-796-848A-23	Sequence 23, Appl
26	50.0	9	9	US-09-782-672-53	Sequence 53, Appl
27	50.0	9	9	US-10-144-644-12	Sequence 12, Appl
28	50.0	9	10	US-09-910-059-28	Sequence 28, Appl
29	47.9	9	9	US-10-161-145-20	Sequence 20, Appl
30	47.9	9	10	US-09-808-037-18	Sequence 18, Appl
31	47.9	9	10	US-09-809-739-4	Sequence 4, Appli
32	47.9	9	10	US-09-974-449-50	Sequence 50, Appl
33	45.8	9	9	US-09-423-800-61	Sequence 61, Appl
34	45.8	9	9	US-09-875-221A-6	Sequence 4, Appli
35	45.8	9	9	US-10-015-535-4	Sequence 5, Appli
36	45.8	9	9	US-09-269-921-5	Sequence 61, Appl
37	45.8	9	9	US-10-182-018-61	Sequence 13, Appl
38	45.8	9	10	US-09-286-240-19	Sequence 19, Appl
39	45.8	9	10	US-09-954-166-13	Sequence 6, Appli
40	45.8	9	10	US-09-949-559-6	Sequence 1, Appli
41	45.8	9	10	US-09-434-965-1	Sequence 13, Appl
42	43.8	7	10	US-09-870-472-13	Sequence 32, Appl
43	43.8	8	9	US-10-046-801-32	Sequence 88, Appl
44	43.8	9	9	US-09-900-590-88	Sequence 26, Appl
45	43.8	9	9	US-09-782-672-26	

ALIGNMENTS

RESULT 1  
US-09-976-787-6  
; Sequence 6, Application US/09976787  
; Patent No. US20020064528A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Zhenping  
; APPLICANT: Witte, Larry  
; TITLE OF INVENTION: Antibodies Specific to KDR and Uses Thereof  
; FILE REFERENCE: 11245/46505  
; CURRENT APPLICATION NUMBER: US/09/976,787  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 09/493,539  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,726  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 6  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-976-787-6  
  
Query Match 64.6%; Score 31; DB 10; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2.7e+05;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 QQLVEYPT 9  
DB 1 QQRSSYPPT 9  
  
RESULT 2  
US-09-865-198-6  
; Sequence 6, Application US/09865198  
; Patent No. US20020103345A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Zhenping  
; TITLE OF INVENTION: Bispecific Immunoglobulin-Like Antigen Binding Proteins and Methc  
; TITLE OF INVENTION: Production  
; FILE REFERENCE: 11245/47102  
; CURRENT APPLICATION NUMBER: US/09/865,198

; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: US 60/206,749  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 6  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-865-198-6

Query Match 64.6%; Score 31; DB 10; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2.7e+05;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
|||  
Db 1 QQRSSYPFT 9  
|||

## RESULT 3

US-09-828-708-51  
; Sequence 51, Application US/09828708  
; Patent No. US20020146753A1  
; GENERAL INFORMATION:  
; APPLICANT: Ditzel, H.  
; APPLICANT: Burton, D.  
; APPLICANT: Schaller, M.  
; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their partic  
; TITLE OF INVENTION: autoimmune disease  
; FILE REFERENCE: 1361,005US1  
; CURRENT APPLICATION NUMBER: US/09/828,708  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 4.70  
; SEQ ID NO 51  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-828-708-51

Query Match 62.5%; Score 30; DB 10; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2.7e+05;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
|||  
Db 1 QQLNSYPLT 9  
|||

## RESULT 4

US-10-091-236-16  
; Sequence 16, Application US/10091236  
; Patent No. US20020168360A1  
; GENERAL INFORMATION:  
; APPLICANT: DINGIVAN, CHRISTINE A.  
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE  
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING INTEGRIN ALPHA-V-BETA-3 ANTAGONISTS  
; TITLE OF INVENTION: COMBINATION WITH OTHER PROPHYLACTIC OR THERAPEUTIC AGENTS  
; FILE REFERENCE: 10271-053-999  
; CURRENT APPLICATION NUMBER: US/10/091,236  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/273,098  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: US 60/316,321  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-10-091-236-16

Query Match 56.2%; Score 27; DB 9; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 QLVVEYPFT 9  
|||  
Db 2 QFTHYPFT 9  
|||

## RESULT 5

US-10-032-482-14  
; Sequence 14, Application US/10032482  
; Publication No. US20020197270A1  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Irun  
; APPLICANT: ROTTER, Varda  
; APPLICANT: Wolkowicz, Roland  
; APPLICANT: Ruiz, Pedro  
; APPLICANT: EREZ-ALON, Neta  
; APPLICANT: HERKEL, Johannes  
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS FOR INDUCTION OF ANTI-TUMOR  
; TITLE OF INVENTION: IMMUNITY  
; FILE REFERENCE: COHEN42  
; CURRENT APPLICATION NUMBER: US/10/032,482  
; CURRENT FILING DATE: 2002-01-02  
; PRIOR APPLICATION NUMBER: US/09/445,602  
; PRIOR FILING DATE: 2001-01-24  
; PRIOR APPLICATION NUMBER: PCT/IL98/00266  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: IL 121041  
; PRIOR FILING DATE: 1997-06-09  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 14  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-032-482-14

Query Match 56.2%; Score 27; DB 9; Length 9;  
Best Local Similarity 55.6%; Pred. No. 2.7e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
|||  
Db 1 QQRSSFPFT 9  
|||

## RESULT 6

US-10-091-313-6  
; Sequence 6, Application US/10091313  
; Publication No. US20030044406A1  
; GENERAL INFORMATION:  
; APPLICANT: DINGIVAN, CHRISTINE  
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE  
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING CD2 ANTAGONISTS IN COMBINATION WITH G  
; TITLE OF INVENTION: PROPHYLACTIC OR THERAPEUTIC AGENTS  
; FILE REFERENCE: 10271-063  
; CURRENT APPLICATION NUMBER: US/10/091,313  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/273,098  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: US 60/346,918  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-10-091-313-6

Query Match 56.2%; Score 27; DB 9; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 QLVVEPFT 9  
| | | | |  
DB 2 QFTHYPFT 9

## RESULT 7

US-09-155-106-6  
; Sequence 6, Application US/09155106  
; Publication No. US20030054003A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: HUMANIZATION OF AN ANTI-CARCINOMABRYONIC  
; TITLE OF INVENTION: ANTIGEN ANTI-IDIOTYPE ANTIBODY AND USE AS A TUMOR VACCINE  
; TITLE OF INVENTION: AND FOR TARGETING APPLICATIONS  
; NUMBER OF SEQUENCES: 45  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/155.106  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/04696  
; FILING DATE: 19-MAR-1997  
; APPLICATION NUMBER: US 60/013,708  
; FILING DATE: 20-MAR-1996  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-155-106-6

Query Match 56.2%; Score 27; DB 9; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 EYPFT 9  
| | | | |  
DB 5 EYPFT 9

## RESULT 8

US-10-091-268-6  
; Sequence 6, Application US/10091268  
; Publication No. US20030068320A1  
; GENERAL INFORMATION:  
; APPLICANT: DINGIVAN, CHRISTINE A  
; TITLE OF INVENTION: METHODS OF ADMINISTERING/DOSING CD2 ANTAGONISTS FOR THE PREVENTION  
; TITLE OF INVENTION: AND TREATMENT OF AUTOIMMUNE DISORDERS OR INFLAMMATORY DISORDERS  
; FILE REFERENCE: 10271-054-999  
; CURRENT APPLICATION NUMBER: US/10/091,268  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/273,098  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: US 60/346,918  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-10-091-268-6

Query Match 56.2%; Score 27; DB 9; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 QLVVEPFT 9  
| | | | |  
DB 2 QFTHYPFT 9

## RESULT 9

US-09-293-854-7  
; Sequence 7, Application US/09293854  
; Patent No. US20020168357A1  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Hing C.  
; Jiao, Jin-an  
; Esperanza, Nieves  
; Lawrence, Luepschen  
; TITLE OF INVENTION: ANTIBODIES FOR INHIBITING BLOOD  
; TITLE OF INVENTION: COAGULATION AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/293,854  
; FILING DATE: 16-Apr-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/814,806  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Corless, Peter F  
; REGISTRATION NUMBER: 33,860  
; REFERENCE/DOCKET NUMBER: 46943  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-293-854-7

Query Match 54.2%; Score 26; DB 9; Length 9;  
Best Local Similarity 55.6%; Pred. No. 2.7e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEPFT 9  
| | | | |  
DB 1 QQVVSPPFT 9

## RESULT 10

US-09-996-288-16

; Sequence 16, Application US/09996288  
; Patent No. US2002017126A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, James  
; APPLICANT: Scott, Koenig  
; APPLICANT: Leslie, Johnson  
; TITLE OF INVENTION: Methods of Administering/Dosing Anti-RSV Antibodies for Prophylaxis  
; FILE OF INVENTION: and treatment  
; FILE REFERENCE: 10271-047-999  
; CURRENT APPLICATION NUMBER: US/09/996,288  
; CURRENT FILING DATE: 2001-11-28  
; NUMBER OF SEQ ID NOS: 259  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-996-288-16

Query Match 54.2%; Score 26; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 2.7e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 QLVVEYPFT 9  
Db 2 QFSGYPFT 9

## RESULT 11

US-09-977-797A-62  
; Sequence 62, Application US/09977797A  
; Publication No. US20030044772A1  
; GENERAL INFORMATION:  
; APPLICANT: Watkins, Jeffrey D.  
; APPLICANT: Huse, William D.  
; TITLE OF INVENTION: Methods for Identifying Ligand Specific Binding Molecules  
; FILE REFERENCE: AME-06805  
; CURRENT APPLICATION NUMBER: US/09/977,797A  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 09/129,026  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 08/905,825  
; PRIOR FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 136  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 62  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-797A-62

Query Match 54.2%; Score 26; DB 9; Length 9;  
Best Local Similarity 55.6%; Pred. No. 2.7e+05;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
Db 1 QQYSRYPLT 9

## RESULT 12

US-09-977-797A-66  
; Sequence 66, Application US/09977797A  
; Publication No. US20030044772A1  
; GENERAL INFORMATION:  
; APPLICANT: Watkins, Jeffrey D.  
; APPLICANT: Huse, William D.  
; APPLICANT: Wu, Herren  
; TITLE OF INVENTION: Methods for Identifying Ligand Specific Binding Molecules  
; FILE REFERENCE: AME-06805  
; CURRENT APPLICATION NUMBER: US/09/977,797A  
; CURRENT FILING DATE: 2002-06-25

; PRIOR APPLICATION NUMBER: 09/129,026  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 08/905,825  
; PRIOR FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 136  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 66  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-797A-66

Query Match 54.2%; Score 26; DB 9; Length 9;  
Best Local Similarity 55.6%; Pred. No. 2.7e+05;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
Db 1 QQYSRYPLT 9

## RESULT 13

US-09-808-037-17  
; Sequence 17, Application US/09808037  
; Patent No. US20020052311A1  
; GENERAL INFORMATION:  
; APPLICANT: HANAN, Eilat  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF  
; FILE REFERENCE: SOLOMON=2D  
; CURRENT APPLICATION NUMBER: US/09/808,037  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 09/629,971  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 09/473,653  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: US 60/152,417  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-09-808-037-17

Query Match 54.2%; Score 26; DB 10; Length 9;  
Best Local Similarity 62.5%; Pred. No. 2.7e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 QLVVEYPFT 9  
Db 2 QRSSYPT 9

## RESULT 14

US-09-808-037-19  
; Sequence 19, Application US/09808037  
; Patent No. US20020052311A1  
; GENERAL INFORMATION:  
; APPLICANT: SOLOMON, Beka  
; APPLICANT: HANAN, Eilat  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF  
; FILE REFERENCE: SOLOMON=2D  
; CURRENT APPLICATION NUMBER: US/09/808,037  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 09/629,971  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 09/473,653

; PRIOR FILING DATE: 1999-12-29  
 ; PRIOR APPLICATION NUMBER: US 60/152,417  
 ; PRIOR FILING DATE: 1999-09-03  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 19  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetic peptide  
 US-09-808-037-19

Query Match 54.2%; Score 26; DB 10; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 2.7e+05;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 QLVYYPFT 9  
 Db 2 QRSYPFT 9

RESULT 15  
 US-09-994-078-2  
 ; Sequence 2, Application US/09994078  
 ; Publication No. US20030032774A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BROWN, WILLIAM  
 ; APPLICANT: DIMAIO, JOHN  
 ; APPLICANT: SCHILLER, PETER  
 ; APPLICANT: MARTEL, RENE  
 ; APPLICANT: MARTEL, FRANCIS  
 ; APPLICANT: MARTEL, DIANE  
 ; APPLICANT: MARTEL, PIERRE  
 ; TITLE OF INVENTION: NOVEL OPIOID PEPTIDES FOR THE TREATMENT OF PAIN  
 ; FILE REFERENCE: MAS/81823/282437  
 ; CURRENT APPLICATION NUMBER: US/09/994,078  
 ; CURRENT FILING DATE: 2001-11-26  
 ; PRIOR APPLICATION NUMBER: 09/159,518  
 ; PRIOR FILING DATE: 1998-09-24  
 ; PRIOR APPLICATION NUMBER: 08/392,918  
 ; PRIOR FILING DATE: 1995-03-03  
 ; PRIOR APPLICATION NUMBER: 08/718,585  
 ; PRIOR FILING DATE: 1996-10-02  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 4  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: peptide  
 US-09-994-078-2

Query Match 52.1%; Score 25; DB 9; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YPFT 9  
 Db 1 YPFT 4

Search completed: April 22, 2003, 13:39:11  
 Job time : 42 secs



GenCore version 5.1.4.p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:29:16 Search time 14 Seconds  
(without alignments)  
18.915 Million cell updates/sec

Title: US-09-674-716B-7  
Perfect score: 48  
Sequence: 1 QQLVEYPPT 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 66399

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	91.7	9	2	US-08-672-345C-27
2	44	91.7	9	2	US-08-214-095D-27
3	36	75.0	9	4	US-08-672-345C-21
4	36	75.0	9	4	US-08-214-095D-21
5	34	70.8	9	4	US-09-406-532-20
6	33	68.8	9	1	US-08-438-123-3
7	32	66.7	9	1	US-08-264-093-26
8	30	62.5	9	2	US-08-672-345C-24
9	30	62.5	9	2	US-08-672-345C-30
10	30	62.5	9	2	US-09-214-095D-24
11	30	62.5	9	4	US-09-214-095D-30
12	28	58.3	9	2	US-08-116-778E-11
13	28	58.3	9	2	US-08-438-562-11
14	28	58.3	9	2	US-08-483-528B-99
15	28	58.3	9	3	US-08-974-899-15
16	28	58.3	9	4	US-09-042-353-368
17	28	58.3	9	4	US-08-758-417A-216
18	28	58.3	9	4	US-09-393-385B-110
19	27	56.2	8	4	US-08-444-818-362
20	27	56.2	8	4	US-08-444-818-363
21	27	56.2	8	4	US-08-444-818-364
22	27	56.2	8	2	US-08-672-345C-81
23	27	56.2	9	4	US-09-214-095D-81
24	27	56.2	9	4	US-09-170-769A-24
25	26	54.2	7	2	US-08-627-173-7
26	26	54.2	7	2	US-08-535-882A-7
27	26	54.2	7	3	US-08-981-384-1

28	26	54.2	7	3	US-09-005-546-7	Sequence 7, Appli
29	26	54.2	8	2	US-08-627-173-6	Sequence 6, Appli
30	26	54.2	8	2	US-08-535-882A-6	Sequence 6, Appli
31	26	54.2	8	3	US-09-005-546-6	Sequence 6, Appli
32	26	54.2	9	2	US-08-627-173-5	Sequence 5, Appli
33	26	54.2	9	2	US-08-535-882A-5	Sequence 5, Appli
34	26	54.2	9	3	US-08-814-806-7	Sequence 7, Appli
35	26	54.2	9	3	US-09-005-546-5	Sequence 5, Appli
36	26	54.2	9	4	US-08-918-148-24	Sequence 24, Appli
37	25	52.1	8	2	US-08-350-260A-521	Sequence 521, App
38	25	52.1	8	2	US-08-350-260A-47	Sequence 47, Appli
39	25	52.1	9	1	US-08-467-420A-48	Sequence 48, Appli
40	25	52.1	9	1	US-08-467-420A-48	Sequence 48, Appli
41	25	52.1	9	1	US-08-470-110A-47	Sequence 47, Appli
42	25	52.1	9	1	US-08-470-110A-48	Sequence 48, Appli
43	25	52.1	9	1	US-08-667-769A-47	Sequence 47, Appli
44	25	52.1	9	1	US-08-667-769A-48	Sequence 48, Appli
45	25	52.1	9	2	US-08-940-371-47	Sequence 47, Appli

## ALIGNMENTS

RESULT 1  
US-08-672-345C-27  
; Sequence 27, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-27

Query Match 91.7%; Score 44; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. NO. 1.9e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 QQLVEYPPT 9  
Db 1 QQLVEYPPT 9

RESULT 2

```
US-09-214-095D-27
; Sequence 27, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 27
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-27

Query Match          91.7%; Score 44; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.9e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9
Db 1 QQFVDYPFT 9

RESULT 3
US-08-672-345C-21
; Sequence 21, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-21

Query Match          75.0%; Score 36; DB 2; Length 9;
Best Local Similarity 66.9%; Pred. No. 1.9e+05;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9
Db 1 QQFVDYPFT 9

us-09-674-716b-7.closed.ra1

US-09-214-095D-21
; Sequence 21, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 21
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-21

Query Match          75.0%; Score 36; DB 4; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.9e+05;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9
Db 1 QQFVDYPFT 9

RESULT 4
US-09-406-532-20
; Sequence 20, Application US/09406532A
; Patent No. 6365154
; GENERAL INFORMATION:
; APPLICANT: Connie L. Erickson-Miller
; APPLICANT: Stephen D. Holmes
; APPLICANT: James D. Winkler
; TITLE OF INVENTION: TIE2 Agonist Antibodies
; FILE REFERENCE: P50843
; CURRENT APPLICATION NUMBER: US/09/406,532A
; CURRENT FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 60/102,098
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)...(9)
; OTHER INFORMATION: light chain CDR 3
US-09-406-532-20

Query Match          70.8%; Score 34; DB 4; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.9e+05;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLVEYPFT 9
Db 2 QRLVEYPFT 9

RESULT 5
US-08-438-123-3
; Sequence 3, Application US/08438123
; Patent No. 5552293
; GENERAL INFORMATION:
; APPLICANT: Lindholm et al
; TITLE OF INVENTION: TUMOR ANTIGEN SPECIFIC ANTIBODY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
US-08-438-123-3

Query Match          75.0%; Score 36; DB 2; Length 9;
Best Local Similarity 66.9%; Pred. No. 1.9e+05;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9
Db 1 QQFVDYPFT 9
```



ADDRESSEE: Lowe, Price, Leblanc & Becker  
STREET: Suite 300, 99 Canal Center Plaza  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC Compatible  
SOFTWARE: DOS Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,123  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/906,350  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: J.G. Mullins  
REGISTRATION NUMBER: 33073  
REFERENCE/DOCKET NUMBER: 149-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 684 1111  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Polypeptide  
US-08-438-123-3

Query Match 68.8%; Score 33; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.9e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLVVEYPT 9  
| : : : : :  
Db 2 QHLEVPPT 9

RESULT 7  
US-08-264-093-26  
Sequence 26, Application US/08264093  
Patent No. 5639863  
GENERAL INFORMATION:  
APPLICANT: Michael D. Dan  
TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES SPECIFIC TO  
TITLE OF INVENTION: CELL CYCLE-INDEPENDENT GLIOMA SURFACE  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ridout & Maybee  
STREET: 2300 Richmond-Adelaide Centre  
STREET: 101 Richmond Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 2U7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette - 3.5 inch, 1.4 Mb storage  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS 6.00  
SOFTWARE: ASCII Editor  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/264,093  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA: No. 5639863 applicable  
ATTORNEY/AGENT INFORMATION:  
NAME: Lake, James R.  
REGISTRATION NUMBER: 31081

REFERENCE/DOCKET NUMBER: NOVOP/106A/7551  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 868-1482  
TELEFAX: (416) 362-0823  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: not applicable  
TOPOLOGY: linear  
US-08-264-093-26

Query Match 66.7%; Score 32; DB 1; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.9e+05;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLVVEYPT 9  
| : : : : :  
Db 2 QRIEFPPT 9

RESULT 8  
US-08-672-345C-24  
Sequence 24, Application US/08672345C  
Patent No. 5948658  
GENERAL INFORMATION:  
APPLICANT: Landry Donald, W.  
TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
NUMBER OF SEQUENCES: 108  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper and Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,345C  
FILING DATE: 24-JUN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/51400  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-672-345C-24

Query Match 62.5%; Score 30; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPT 9  
| : : : : :  
Db 1 QHFEYPT 9

RESULT 9  
US-08-672-345C-30  
Sequence 30, Application US/08672345C

Patent No. 5948658  
GENERAL INFORMATION:  
APPLICANT: Landry, Donald, W.  
TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
NUMBER OF SEQUENCES: 108  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper and Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,345C  
FILING DATE: 24-JUN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/51400  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-672-345C-30

Query Match 62.5%; Score 30; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QQLVEYPT 9  
Db 1 QHFEDYPT 9

RESULT 10  
US-09-214-095D-24  
Sequence 24, Application US/09214095D  
Patent No. 6280987  
GENERAL INFORMATION:  
APPLICANT: Landry, Donald  
TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
FILE REFERENCE: 51400-A-PCT-US  
CURRENT APPLICATION NUMBER: US/09/214,095D  
CURRENT FILING DATE: 1999-07-19  
NUMBER OF SEQ ID NOS: 121  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 24  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Murinae gen. sp.  
US-09-214-095D-24

Query Match 62.5%; Score 30; DB 4; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QQLVEYPT 9  
Db 1 QHFEDYPT 9

RESULT 11  
US-09-214-095D-30  
Sequence 30, Application US/09214095D  
Patent No. 6280987  
GENERAL INFORMATION:  
APPLICANT: Landry, Donald  
TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
FILE REFERENCE: 51400-A-PCT-US  
CURRENT APPLICATION NUMBER: US/09/214,095D  
CURRENT FILING DATE: 1999-07-19  
NUMBER OF SEQ ID NOS: 121  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 30  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Murinae gen. sp.  
US-09-214-095D-30

Query Match 62.5%; Score 30; DB 4; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QQLVEYPT 9  
Db 1 QHFEDYPT 9

RESULT 12  
US-08-116-778E-11  
Sequence 11, Application US/08116778E  
Patent No. 5830470  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, KAZUYASU  
APPLICANT: KOIKE, MASAMICHI  
APPLICANT: SHITARA, KENYA  
APPLICANT: HANAI, NOBUO  
APPLICANT: KUWANA, YOSHIHISA  
APPLICANT: HASEGAWA, MAMORU  
TITLE OF INVENTION: HUMANIZED ANTIBODIES  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/116,778E  
FILING DATE: 07-SEP-93  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 249-59  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)816-4000  
TELEFAX: (703)816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-116-778E-11

Query Match 58.3%; Score 28; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
|||  
Db 1 QQRSSYPFT 9

RESULT 13  
US-08-438-562-11  
; Sequence 11, Application US/08438562  
; Patent No. 5874255  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, KAZUYASU  
; APPLICANT: KOIKE, MASAMICHI  
; APPLICANT: SHITARA, KENYA  
; APPLICANT: HANAI, NOBUO  
; APPLICANT: KIWANA, YOSHIIISA  
; APPLICANT: HASEGAWA, MAMORU  
; TITLE OF INVENTION: HUMANIZED ANTIBODIES  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/438,562  
; FILING DATE: 10-MAY-95  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/116,778  
; FILING DATE: 07-SEP-93  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 249-76  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)816-4000  
; TELEFAX: (703)816-4100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-438-562-11

Query Match 58.3%; Score 28; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
|||  
Db 1 QQRSSYPFT 9

RESULT 14  
US-08-483-528B-99  
; Sequence 99, Application US/08483528B  
; Patent No. 5939532  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, KAZUYASU  
; APPLICANT: KOIKE, MASAMICHI  
; APPLICANT: SHITARA, KENYA  
; APPLICANT: HANAI, NOBUO

; APPLICANT: KUWANA, YOSHIIISA  
; APPLICANT: HASEGAWA, MAMORU  
; TITLE OF INVENTION: HUMANIZED ANTIBODIES  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,528B  
; FILING DATE: 07-JUN-95  
; CLASSIFICATION: 536  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)816-4000  
; TELEFAX: (703)816-4100  
; INFORMATION FOR SEQ ID NO: 99:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acids  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-483-528B-99

Query Match 58.3%; Score 28; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
|||  
Db 1 QQRSSYPFT 9

RESULT 15  
US-08-974-899-15  
; Sequence 15, Application US/08974899  
; Patent No. 6037454  
; GENERAL INFORMATION:  
; APPLICANT: Presta, Leonard G.  
; APPLICANT: Jardieu, Paula M.  
; TITLE OF INVENTION: Humanized Anti-CD11a Antibodies  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/974,899  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/031971  
; FILING DATE: 11/27/96  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P1014R1  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-1994  
 ; TELEFAX: 650/952-9881  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9 amino acids  
 ; TYPE: Amino Acid  
 ; TOPOLOGY: Linear  
 ;  
 US-08-974-899-15

Query Match 58.3%; Score 28; DB 3; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQLVVEYPT 9  
 Db 1 QQHNEYPLT 9

Search completed: April 22, 2003, 13:32:17  
 Job time : 15 secs

GenCore version 5.1.4 p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:38:27 ; Search time 43 Seconds  
(without alignments)  
11.178 Million cell updates/sec

Title: US-09-674-716B-9  
Perfect score: 33  
Sequence: 1 GYWS 5  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 206

Minimum DB seq length: 0  
Maximum DB seq length: 5

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	60.6	5	2 JH0253	gut pentapeptide -
2	16	48.5	5	2 A32516	cholecystokinin-5
3	13	39.4	4	2 S09478	globulin IV alpha
4	13	39.4	5	2 P70308	Ig heavy chain CRD
5	11	33.3	3	3 F37196	bradykinin-potenti
6	11	33.3	4	2 A34626	RPCB-related neuro
7	11	33.3	4	2 B53284	T-cell receptor be
8	11	33.3	4	2 P70661	T-cell receptor be
9	11	33.3	5	2 A60803	neuropeptide - sea
10	11	33.3	5	2 G37196	bradykinin-potenti
11	11	33.3	5	2 P70281	Ig heavy chain CRD
12	11	33.3	5	2 P70729	T-cell receptor be
13	11	33.3	5	2 P70580	T-cell receptor be
14	9	27.3	3	3 I50412	gene p20K protein
15	9	27.3	4	2 A32480	achatin-I - giant
16	9	27.3	5	2 S70154	pap fibrial regul
17	9	27.3	5	2 B37325	pap fibrial regul
18	9	27.3	5	2 A40469	dnazX-like protein
19	9	27.3	5	2 A37114	hypoxanthine phosph
20	9	27.3	5	2 B61445	Leu-enkephalin - b
21	9	27.3	5	2 A61445	Met-enkephalin - b
22	9	27.3	5	2 B61168	cocoonase (EC 3.4.
23	8	24.2	3	3 GRHU	growth-modulating
24	8	24.2	4	2 P01140	carbon-monoxide de
25	8	24.2	4	2 I38888	COI intron 16 prot
26	8	24.2	5	2 P70278	Ig heavy chain CRD
27	8	24.2	5	2 S69237	surface protein te
28	7	21.2	3	3 A22565	R-phycoerythrin al
29	7	21.2	3	3 A43391	TRH-like tripeptid

30	7	21.2	3	3 S68328	blood cell protein
31	7	21.2	4	2 A32039	tyrosine-melanocyt
32	7	21.2	4	2 A37832	phenol 2-monooxyge
33	7	21.2	4	2 I61883	protamine P1 - ora
34	7	21.2	4	2 P70240	Ig mu chain V regi
35	7	21.2	4	2 S43959	Ig mu chain V regi
36	7	21.2	4	2 I37013	protamine P1 - Cer
37	7	21.2	4	2 I84439	protamine P1 - sav
38	7	21.2	5	1 HOROHA	proctolin - Americ
39	7	21.2	5	2 I39964	ribosomal protein
40	7	21.2	5	2 I39966	ribosomal protein
41	7	21.2	5	2 I39965	ribosomal protein
42	7	21.2	5	2 E60274	major protein anti
43	7	21.2	5	2 F22565	R-phycoerythrin ga
44	7	21.2	5	2 P00009	angiotensin-conver
45	7	21.2	5	2 P00689	photosystem I 10.4

## ALIGNMENTS

### RESULT 1

JH0253  
gut pentapeptide - Japanese eel  
C;Species: Anguilla japonica (Japanese eel)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 11-Apr-1995  
C;Accession: JH0253  
R;Uesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.; Ando, M.  
Biochem. Biophys. Res. Commun. 180, 828-832, 1991  
A;Title: Structure and function of a pentapeptide isolated from the gut of the eel.  
A;Reference number: JH0253; MUID:92062113; PMID:1953755  
A;Accession: JH0253  
A;Molecule type: protein  
A;Residues: 1-5 <UNS>  
A;Experimental source: gut  
C;Comment: This peptide increased basal tone of the circular muscle of the esophagogastric, and of the circular muscle of the gastro-intestinal junction.

Query Match 60.6%; Score 20; DB 2; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYW 3  
Db 1 GFW 3

### RESULT 2

A32516  
cholecystokinin-5 - dog  
N;Alternate names: CCK-5  
C;Species: Canis lupus familiaris (dog)  
C;Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 18-Aug-2000  
C;Accession: A32516  
R;Shively, J.; Reeve Jr., J.R.; Eysselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh, J.H.  
Am. J. Physiol. 252, G272-G275, 1987  
A;Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and intest  
A;Reference number: A32516; MUID:87153871; PMID:3826354  
A;Accession: A32516  
A;Molecule type: protein  
A;Residues: 1-5 <SHI>  
C;Comment: This peptide corresponds to the five carboxyl-terminal residues of cholecysto  
C;Superfamily: gastrin  
C;Keywords: amidated carboxyl end; neuropeptide  
F;5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 48.5%; Score 16; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WM 4  
Db 2 WM 3

```

RESULT 3
S09478
Globulin IV alpha subunit gamma chain, seed - cucurbit (fragment)
N:Alternate names: 11S globulin alpha subunit gamma chain
C:Species: Cucurbita sp. (cucurbit)
C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 08-Nov-1996
C:Accession: S09478
R:Ohmura, M.; Hara, I.; Matsubara, H.
Plant Cell Physiol. 21, 157-167, 1980
A:Title: Pumpkin (Cucurbita sp.) seed globulin IV. Terminal sequences of the acidic and
A:Reference number: S09066
A:Accession: S09478
A:Molecule type: protein
A:Residues: 1-4 <OHM>

Query Match 39.4%; Score 13; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GY 2
||
Db 3 GY 4

RESULT 4
PT0308
IG heavy chain CRD3 region (clone 6-88) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0308
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovers, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0308
A:Molecule type: DNA
A:Residues: 1-5 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 39.4%; Score 13; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WMS 5
||
Db 2 WES 4

RESULT 5
F37196
bradykinin-potentiating peptide 6 - island jaxaraca
C:Species: Bothrops insularis (island jaxaraca)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: F37196
R:Cintrà, A.C.O.; Vieira, C.A.; Giglio, J.M.
J. Protein Chem. 9, 221-227, 1990
A:Title: Primary structure and biological activity of bradykinin potentiating peptides f
A:Reference number: A37196; MUID:90351557; PMID:2386615
A:Accession: F37196
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-3 <CIN>
C:Keywords: pyroglutamic acid
P:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 33.3%; Score 11; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3,

```

```

Db 3 W 3

RESULT 6
A34626
RPCH-related neuropeptide - ferruginous spindle
C:Species: Fuscus ferrugineus (ferruginous spindle)
C:Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 31-Dec-1993
C:Accession: A34626
R:Kuroki, Y.; Kanda, T.; Kubota, I.; Fujisawa, Y.; Ikeda, T.; Miura, A.; Minamitake, Y.;
Biochem. Biophys. Res. Commun. 167, 273-279, 1990
A:Title: A molluscan neuropeptide related to the crustacean hormone, RPCH.
A:Reference number: A34626; MUID:90179762; PMID:2310394
A:Accession: A34626
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4 <KUR>
C:Keywords: neuropeptide

Query Match 33.3%; Score 11; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3
||
Db 4 W 4

RESULT 7
B53284
T-cell receptor beta 2 chain D region, Dbeta2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: B53284
R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity and
A:Reference number: A53284; MUID:91342695; PMID:1678859
A:Accession: B53284
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4 <HAP>
A:Cross-references: GB:S60737; NID:9233916; PIDN:AAB19518.1; PID:9233918
A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60738)
C:Keywords: T-cell receptor

Query Match 33.3%; Score 11; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3
||
Db 2 W 2

RESULT 8
PT0661
T-cell receptor beta chain V-D-J region (121-1BV) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0661
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0661
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FEE>
A:Experimental source: day 4 postnatal thymus, strain BALB/c
C:Keywords: T-cell receptor

```

Query Match 33.3%; Score 11; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3  
 |  
 Db 3 W 3

RESULT 9  
 AG0803  
 neuropeptide - sea anemone (Anthopleura elegantissima)  
 C:Species: Anthopleura elegantissima  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
 C:Accession: AG0803  
 R:Graff, D.; Grimmelikhuijzen, C.J.P.  
 Brain Res. 442, 354-358, 1988  
 A>Title: Isolation of <Glu-Ser-Lu-Arg-Trip-NH-2, a novel neuropeptide from sea anemones.  
 A:Reference number: AG0803; MUID:88222764; PMID:2897223  
 A:Accession: AG0803  
 A:Molecule type: protein  
 A:Residues: 1-5 <GSA>  
 C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:5/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 33.3%; Score 11; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3  
 |  
 Db 5 W 5

RESULT 10  
 G37196  
 bradykinin-potentiating peptide 7 - island jararaca  
 C:Species: Bothrops insularis (Island jararaca)  
 C>Date: 14-Feb-1992 #sequence\_revision 01-Dec-1992 #text\_change 05-Aug-1994  
 C:Accession: G37196  
 R:Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.  
 J. Protein Chem. 9, 221-227, 1990  
 A>Title: Primary structure and biological activity of bradykinin potentiating peptides 6  
 A:Reference number: A37196; MUID:90351557; PMID:2386615  
 A:Accession: G37196  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-5 <CIN>  
 C:Keywords: pyroglutamic acid  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 33.3%; Score 11; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3  
 |  
 Db 3 W 3

RESULT 11  
 PT0281  
 Ig heavy chain CRD3 region (clone 4-91C) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0281  
 R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A>Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
 A:Reference number: PT0222; MUID:91108337; PMID:1899102  
 A:Accession: PT0281  
 A:Molecule type: DNA

A:Residues: 1-5 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 33.3%; Score 11; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3  
 |  
 Db 4 W 4

RESULT 12  
 PT0729  
 T-cell receptor beta chain V-D-J region (120-1J) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0640; PT0685; PT0729  
 R:Feeney, A.J.  
 J. Exp. Med. 174, 115-124, 1991  
 A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601; PMID:1711558  
 A:Accession: PT0640  
 A>Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-5 <FEE>  
 A:Experimental source: newborn thymus, strain BALB/c, clone 120-1J  
 A:Accession: PT0685  
 A>Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-5 <FE2>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, clone 154-1C  
 A:Accession: PT0729  
 A>Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-5 <FE3>  
 A:Experimental source: newborn thymus, strain BALB/c, clone 135-1AG  
 C:Keywords: T-cell receptor

Query Match 33.3%; Score 11; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3  
 |  
 Db 5 W 5

RESULT 13  
 PT0580  
 T-cell receptor beta chain V-D-J region (159-2B) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0580  
 R:Feeney, A.J.  
 J. Exp. Med. 174, 115-124, 1991  
 A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601; PMID:1711558  
 A:Accession: PT0580  
 A>Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-5 <FEE>  
 A:Experimental source: day 19 fetal thymus, strain BALB/c  
 C:Keywords: T-cell receptor

Query Match 33.3%; Score 11; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 W 3  
 |  
 Db 4 W 4

## RESULT 14

I50412  
 gene p20K protein - chicken (fragment)  
 C;Species: Gallus gallus (Chicken)  
 C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
 C;Accession: I50412  
 R;Mao, P. L.; Beauchemin, M.; Bedard, P. A.  
 J. Biol. Chem. 268, 8131-8139, 1993  
 A;Title: Quiescence-dependent activation of the p20K promoter in growth-arrested chicken  
 A;Reference number: A46643; MUID:93216790; PMID:8463325  
 A;Accession: I50412  
 A;Status: Preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-3 <MAO>  
 A;Cross-references: GB:L02537; NID:g212616; PID:g212617  
 C;Genetics:  
 A;Gene: p20K

Query Match 27.3%; Score 9; DB 3; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MS 5  
 ||  
 Db 1 MS 2

## RESULT 15

A32480  
 achatin-I - giant African snail  
 N;Contains: achatin-II  
 C;Species: Achatina fulica (giant African snail)  
 C;Date: 12-Oct-1989 #sequence\_revision 12-Oct-1989 #text\_change 17-Mar-1999  
 C;Accession: A32480  
 R;Kamatani, Y.; Minakata, H.; Kenny, P. T. M.; Iwashita, T.; Watanabe, K.; Funase, K.; Sur  
 Biochem. Biophys. Res. Commun. 160, 1015-1020, 1989  
 A;Title: Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina fulica feru  
 A;Reference number: A32480; MUID:89273551; PMID:2597281  
 A;Accession: A32480  
 A;Molecule type: protein  
 A;Residues: 1-4 <KAM>  
 A;Note: stereochemistry of the active form confirmed by chemical synthesis  
 R;Ishida, T.; In, Y.; Inoue, M.; Yasuda-Kamatani, Y.; Minakata, H.; Iwashita, T.; Nomoto  
 FEBS Lett. 307, 253-256, 1992  
 A;Title: Effect of the D-Phe(2) residue on molecular conformation of an endogenous neuro  
 (H-Gly-Phe-Ala-Asp-OH).  
 A;Reference number: A44691; MUID:92354723; PMID:1644179  
 A;Contents: annotation; X-ray crystallography, 0.85 angstroms  
 A;Note: achatin-II has L-phenylalanine  
 C;Keywords: D-amino acid  
 P;2/Modified site: D-phenylalanine (Phe) #status experimental

Query Match 27.3%; Score 9; DB 2; Length 4;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GY 2  
 ||  
 Db 1 GP 2

Search completed: April 22, 2003, 13:42:27  
 Job time : 44 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:32:21 ; Search time 11 Seconds  
(without alignments)  
18.853 Million cell updates/sec

Title: US-09-674-716B-9

Perfect score: 33  
Sequence: 1 GYWS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 38

Minimum DB seq length: 0  
Maximum DB seq length: 5

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	45.5	4	1 OCP3 OCTMI	P58649 octopus min
2	15	45.5	5	1 UF01 MOUSE	P38639 mus musculus
3	11	33.3	5	1 BPP7 BOTIN	P30425 bothrops in
4	10	30.3	5	1 PAP2 PARMA	P81864 pardachirus
5	10	30.3	5	1 RE32 LITRU	P82073 litoria rub
6	9	27.3	4	1 ACHI ACHFU	P35904 achatina fu
7	9	27.3	4	1 OCP1 OCTMI	P58648 octopus min
8	8	24.2	3	1 GRWV HUMAN	P01157 homo sapien
9	8	24.2	4	1 DCML PSECH	P19916 pseudomonas
10	7	21.2	4	1 FAR3 HIRME	P42562 hirudo medi
11	7	21.2	4	1 FAR4 HIRME	P42563 hirudo medi
12	7	21.2	4	1 FYRI ANTEL	P58706 anthopleura
13	7	21.2	5	1 AL14 CARMA	P81817 carcinus ma
14	7	21.2	5	1 FARP ANTR	P41853 artiposthi
15	7	21.2	5	1 PRCT PERAM	P04373 periplaneta
16	7	21.2	5	1 PSK DAUCA	P58261 daucus caro
17	6	18.2	4	1 DCMS PSECH	P19918 pseudomonas
18	6	18.2	4	1 E051 HUMAN	P02731 homo sapien
19	6	18.2	4	1 PMRF MACNI	P01162 macrocallis
20	6	18.2	5	1 B10A CJTER	P13071 citrobacter
21	6	18.2	5	1 B10B CJTER	P12997 citrobacter
22	6	18.2	5	1 TP18 CANFA	P54714 canis famil
23	6	18.2	5	1 UXA4 CHLTR	P38005 chlamydia t
24	5	15.2	4	1 E103 LITRU	P82099 litoria rub
25	4	12.1	4	1 FFKA ANTEL	P58705 anthopleura
26	4	12.1	4	1 RM01 YEAST	P36515 saccharomyc
27	4	12.1	5	1 RE11 LITRU	P82070 litoria rub
28	4	12.1	5	1 RE21 LITRU	P82071 litoria rub
29	4	12.1	5	1 RE31 LITRU	P82072 litoria rub
30	4	12.1	5	1 UC22 MAIZE	P80628 zea mays (m
31	3	9.1	4	1 FLRF HIRME	P42561 hirudo medi
32	3	9.1	4	1 FLRN ANTEL	P58707 anthopleura
33	3	9.1	5	1 E104 LITRU	P82100 litoria rub

34 3 9.1 5 1 SUGA ACHDO P19991 acheta dome  
35 2 6.1 3 1 THYL PIG P01151 sus scrofa  
36 1 3.0 3 1 LUXE VIBFI P24272 vibrio fusc  
37 1 3.0 4 1 TUFT HUMAN P01858 homo sapien  
38 1 3.0 5 1 TRM3\_ECOLI P13973 escherichia

ALIGNMENTS

RESULT 1  
OCP3 OCTMI  
ID OCP3 OCTMI STANDARD; PRT; 4 AA.  
AC P58649;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Cardioactive peptides Ocp-3/Ocp-4.  
OS Octopus minor (Octopus).  
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
OC Incirrata; Octopodidae; Octopus.  
OX NCBI\_TaxID=89766;  
RN [1]  
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
RC TISSUE=Brain;  
RX MEDLINE=20336815; PubMed=10876044;  
RA Iwakoshi E., Hisada M., Minakata H.;  
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
RT Octopus minor.";  
RL Peptides 21:623-630(2000).  
CC -!- FUNCTION: Cardioactive; has both positive chronotropic and  
CC inotropic effects on the heart. Ocp-4 is a 1000 time less  
CC active than Ocp-3.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- PTM: Ocp-4 has D-Ser instead of L-Ser.  
CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.  
KW Hormone; D-amino acid.  
FT MOD RES 2 2 D-SERINE (IN OCP-4).  
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 45.5%; Score 15; DB 1; Length 4;  
Best Local Similarity 66.7%; Fred. No. 1.1e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GYW 3  
| | |  
DB 1 GSW 3

RESULT 2  
UF01 MOUSE  
ID UF01 MOUSE STANDARD; PRT; 5 AA.  
AC P38639;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Fibroblast;  
RX MEDLINE=95009907; PubMed=7523108;  
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;  
RT "Separation and sequencing of familial and novel murine proteins  
RT using preparative two-dimensional gel electrophoresis.";  
RL Electrophoresis 15:735-745(1994).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 6.6, ITS MW IS: 19 kDa.  
FT NON TER 5 5  
SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

```

Query Match      45.5%; Score 15; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYW 3
Db 3 GRW 5

RESULT 3
BPP7_BOTIN
ID BPP7_BOTIN STANDARD; PRT; 5 AA.
AC P30425;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Bradykinin-potentiating peptide S5,2 (5A) (Angiotensin-converting
enzyme inhibitor).
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidogaulia; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
DR PIR; G37196; G37196.
KW Hypotensive agent; Venom.
FT MOD RES 1
SQ SEQUENCE 5 AA; 629 MW; 776DC37326B00000 CRC64;

Query Match      33.3%; Score 11; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 W 3
Db 3 W 3

RESULT 4
PAP2_PARMA
ID PAP2_PARMA STANDARD; PRT; 5 AA.
AC P81864;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pardaxin II (PXII) (Fragment).
OS Pardachirus marmoratus (Red sea mosses sole).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuroctetiformes;
OC Soleioidae; Soleidae; Pardachirus.
OX NCBI_TaxID=31087;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=87057369; PubMed=3782138;
RA Lazarovici P., Primor N., Loew L.M.;
RT "Purification and pore-forming activity of two hydrophobic
peptides from the secretion of the Red sea mosses sole (Pardachirus
marmoratus).";
RJ J. Biol. Chem. 261:16704-16713(1986).

CC -1- FUNCTION: EXHIBITS UNUSUAL SHARK REPELLENT AND SURFACTANT
PROPERTIES. FORMS VOLTAGE-DEPENDENT, ION-PERMEABLE CHANNELS
IN MEMBRANES. AT HIGH CONCENTRATION CAUSES CELL MEMBRANE LYSIS.
CC -1- SUBUNIT: MONOMER. IN AQUEOUS SOLUTION EXISTS AS A TETRAMER.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.
KW Toxin.
FT NON_TER 5
SQ SEQUENCE 5 AA; 614 MW; 7769C9C8100000 CRC64;

Query Match      30.3%; Score 10; DB 1; Length 5;
Best Local Similarity 33.3%; Pred. No. 1.1e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYW 3
Db 1 GFF 3

RESULT 5
RE32_LITRU
ID RE32_LITRU STANDARD; PRT; 5 AA.
AC P82073;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Rubellidin 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
Litoria electrica. Comparison with the skin peptides from Litoria
rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9C862A00000 CRC64;

Query Match      30.3%; Score 10; DB 1; Length 5;
Best Local Similarity 33.3%; Pred. No. 1.1e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYW 3
Db 2 GFF 4

RESULT 6
ACH1_ACHFU
ID ACH1_ACHFU STANDARD; PRT; 4 AA.
AC P35904;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Achatin-I.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Achatinacea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RX STRAIN=Ferussac; TISSUE=Ganglion;
RX MEDLINE=89273551; PubMed=2597281;
RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
Funase K., Sun X.P., Yonggiri A., Kim K.H., Novales-Li P.,

```

RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;  
 RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina  
 RL fulica Ferussac containing a D-amino acid residue";  
 RN Biochem. Biophys. Res. Commun. 160:1015-1020(1989).  
 (2)  
 RP CHARACTERIZATION.  
 RC STRAIN=Ferussac; TISSUE=Heart atrium;  
 RX MEDLINE=91264856; PubMed=1675568;  
 RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,  
 RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;  
 RT "Purification of achatin-I from the atria of the African giant snail,  
 RT Achatina fulica, and its possible function";  
 RL Biochem. Biophys. Res. Commun. 177:847-853(1991).  
 RN (3)  
 RP X-RAY CRYSTALLOGRAPHY.  
 RX MEDLINE=93014529; PubMed=1399265;  
 RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,  
 RA Iwashita T., Nomoto K.;  
 RT "Crystal structure and molecular conformation of achatin-I  
 RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a  
 D-amino acid residue";  
 RL Int. J. Pept. Protein Res. 39:258-264(1992).  
 CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY  
 CC AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY  
 CC NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE  
 CC HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.  
 DR PIR; A32480; A32480.  
 KW Hormone; D-amino acid.  
 FT MOD RES 2 2 D-PHENYLALANINE.  
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;  
 Query Match 27.3%; Score 9; DB 1; Length 4;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GY 2  
 DB 1 GF 2  
 RESULT 7  
 OCPI OCTMI STANDARD; PRT; 4 AA.  
 AC P58648;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cardioactive peptides Ocp-1/Ocp-2.  
 OS Octopus minor (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 OC Incirrata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=89766;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
 RC TISSUE=Brain;  
 RX MEDLINE=20336815; PubMed=10876044;  
 RA Iwakoshi E., Hisada M., Minakata H.;  
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
 RT Octopus minor";  
 RL Peptides 21:623-630(2000).  
 CC -1- FUNCTION: Cardioactive; has both positive chronotropic and  
 CC inotropic effects on the heart. Ocp-2 is a 1000 time less  
 CC active than Ocp-1.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PTM: Ocp-2 has L-Phe instead of D-Phe.  
 CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=WALDI.  
 KW Hormone; D-amino acid.  
 FT MOD RES 2 2 D-PHENYLALANINE.  
 SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;  
 Query Match 27.3%; Score 9; DB 1; Length 4;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GY 2  
 DB 1 GF 2

## RESULT 8

GRWM HUMAN STANDARD; PRT; 3 AA.  
 ID GRWM HUMAN  
 AC P01157;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last annotation update)  
 DE Growth-modulating peptide.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77162369; PubMed=858356;  
 RA Schlesinger D.H., Pickart L., Thaler M.M.;  
 RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine";  
 RL Experientia 33:324-325(1977).  
 CC -1- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE  
 CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.  
 DR PIR; A01421; GKHU.  
 SQ SEQUENCE 3 AA; 340 MW; 6331E81000000000 CRC64;

Query Match 24.2%; Score 8; DB 1; Length 3;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GY 2  
 DB 1 GH 2

## RESULT 9

DCML\_PSECH STANDARD; PRT; 4 AA.  
 AC P19916;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO  
 DE dehydrogenase subunit L) (CO-DH L) (Fragment).  
 GN CUTL.  
 OS Pseudomonas carboxydohydrogena.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Bradyrhizobium group.  
 OX NCBI\_TaxID=290;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=90055678; PubMed=281128;  
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;  
 RT "Homology and distribution of CO dehydrogenase structural genes in  
 RT carboxydrotrophic bacteria";  
 RL Arch. Microbiol. 152:335-341(1989).  
 CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon  
 CC dioxide.  
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced  
 CC acceptor.  
 CC -1- COFACTOR: Molybdenum (molybdopterin).  
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND  
 CC SMALL.  
 DR PIR; PLO140; PLO140.  
 KW Oxidoreductase; Molybdenum.  
 FT NON TER 4 4  
 SQ SEQUENCE 4 AA; 441 MW; 7761E876F0000000 CRC64;

Query Match 24.2%; Score 8; DB 1; Length 4;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;

Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GY 2

DB 2 GH 3

# RESULT 10

FAR3\_HIRME

ID FAR3\_HIRME

AC P42562;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE FMRFamide-like neuropeptide YMRP-amide.

OS Hirudo medicinalis (Medicinal leech).

OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;

OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.

OX NCBI\_TaxID=6421;

RN [1]

RP SEQUENCE.

RX MEDLINE=92195954; PubMed=1686933;

RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;

RT "Identification of RFamide neuropeptides in the medicinal leech.;"

RL Peptides 12:897-908(1991).

CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)

CC FAMILY.

KW Neuropeptide; Amidation.

FT MOD\_RES 4

SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

## Query Match

Best Local Similarity 100.0%; Pred. No. 1.1e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 Y 2

DB 1 Y 1

# RESULT 11

FAR4\_HIRME

ID FAR4\_HIRME

AC P42563;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE FMRFamide-like neuropeptide YMRP-amide.

OS Hirudo medicinalis (Medicinal leech).

OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;

OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.

OX NCBI\_TaxID=6421;

RN [1]

RP SEQUENCE.

RX MEDLINE=92195954; PubMed=1686933;

RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;

RT "Identification of RFamide neuropeptides in the medicinal leech.;"

RL Peptides 12:897-908(1991).

CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)

CC FAMILY.

KW Neuropeptide; Amidation.

FT MOD\_RES 4

SQ SEQUENCE 4 AA; 616 MW; 69D4068B30000000 CRC64;

## Query Match

Best Local Similarity 100.0%; Pred. No. 1.1e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 Y 2

DB 1 Y 1

# RESULT 12

FYRI\_ANTEL

ID FYRI\_ANTEL

AC P58706;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Antho-Riamide I [Contains: Antho-Riamide II].

OS Anthopleura elegantissima (Sea anemone).

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;

OC Nynanthaeae; Actiniidae; Anthopleura.

OX NCBI\_TaxID=6110;

RN [1]

RP SEQUENCE.

RX MEDLINE=92270459; PubMed=1821096;

RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.;

RT "Isolation of two novel neuropeptides from sea anemones: the unusual,

biologically active L-3-phenylalanyl-tyr-Arg-Ile-NH2 and its

des-phenylalanyl fragment Tyr-Arg-Ile-NH2.;"

RL Peptides 12:1165-1173(1991).

RN [2]

RP FUNCTION.

RX MEDLINE=93391436; PubMed=8397415;

RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;

RT "The expansion behaviour of sea anemones may be coordinated by two

inhibitory neuropeptides, Antho-Kamide and Antho-Riamide.;"

Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).

CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle

groups. May be involved in the expansion phase of feeding

behaviour in sea anemones.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Neuron-specific.

DR InterPro; IPR01023; Hsp70.

KW Neuropeptide; Amidation.

FT CHAIN 1 4

FT MOD\_RES 2 4

FT MOD\_RES 1 1

FT MOD\_RES 4 4

SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;

## Query Match

Best Local Similarity 100.0%; Pred. No. 1.1e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 Y 2

DB 2 Y 2

# RESULT 13

ALI4\_CARMA

ID ALI4\_CARMA

AC P81817;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE Carcinus maenas (Common shore crab) (Green crab).

OS Carcinus maenas (Common shore crab) (Green crab).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;

OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;

OC Brachyura; Eubrachyura; Portunoidea; Portunidae; Carcinus.

OX NCBI\_TaxID=6759;

RN [1]

RP SEQUENCE.

RX MEDLINE=98121193; PubMed=9461295;

RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.;

RT "Isolation and identification of multiple neuropeptides of the

allatostatin superfamily in the shore crab Carcinus maenas.;"

Eur. J. Biochem. 250:727-734(1997).

CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.

CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 5 5 AMIDATION (POTENTIAL).  
 SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 21.2%; Score 7; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 Y 2

Db 1 Y 1

## RESULT 14

FARP\_ARTTR ID FARP\_ARTTR STANDARD; PRT; 5 AA.

AC P41853;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE FMRamide-like neuropeptide RYRFP-amide.

OS Artiposatha triangulata.

OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;

OC Terricola; Geoplanidae; Arthurdendyus.

OX NCBI\_TaxID=132421;

RN [1]

RP SEQUENCE, AND SYNTHESIS.

RX MEDLINE=94211927; PubMed=7909164;

RA Maule A.G., Shaw C., Halton D.W., Curry W.J., Thim L.;

RT "RYRFPamide: a turbellarian FMRamide-related peptide (FARP).";

RL Regul. Pept. 50:37-43(1994).

CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)

CC FAMILY

KW Neuropeptide; Amidation.

FT MOD\_RES 5 5 AMIDATION.

SQ SEQUENCE 5 AA; 754 MW; 69D4004B4600000 CRC64;

Query Match

21.2%; Score 7; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 Y 2

Db 2 Y 2

## RESULT 15

PRCT\_PERAM

ID PRCT\_PERAM STANDARD; PRT; 5 AA.

AC P01373;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-FEB-1995 (Rel. 31, Last annotation update)

DE Proctolin.

OS Periplaneta americana (American cockroach),

OS Limulus polyphemus (Atlantic horseshoe crab), and

OS Carcinus maenas (Common shore crab) (Green crab).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;

OC Blattodea; Blattidae; Periplaneta.

OX NCBI\_TaxID=6978, 6850, 6759;

RN [1]

RP SEQUENCE.

RC SPECIES=P.americana;

RX MEDLINE=76074708; PubMed=576;

RA Starratt A.N., Brown B.E.;

RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter

in insects";

RL Life Sci. 17:1253-1256(1975).

RN [2]

RP BIOLOGICAL SOURCE.

RC SPECIES=P.americana;

RX MEDLINE=81225865; PubMed=6113690;  
 RA O'Shea M., Adams M.E.;  
 RT "Pentapeptide (proctolin) associated with an identified neuron."; Science 213:567-569(1981).  
 RL [3]

RN SEQUENCE.

RC SPECIES=L.polyphemus;

RX MEDLINE=90287800; PubMed=2356151;

RA Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,

RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,

RA Shabanowitz J.;

RT "Identification of proctolin in the central nervous system of the

horsehoe crab, Limulus polyphemus.";

RL Peptides 11:205-211(1990).

RN [4]

RP SEQUENCE.

RC SPECIES=C.maenas;

RX MEDLINE=86232789; PubMed=2872661;

RA Stangler J., Dirksen H., Keller R.;

RT "Identification and immunocytochemical localization of proctolin in

pericardial organs of the shore crab, Carcinus maenas.";

RL Peptides 7:67-72(1986).

CC -1- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY,

MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.

CC -1- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN

THE CRAB PERICARDIAL ORGANS.

DR PIR; A01644; HOROHA.

DR PIR; A60411; A60411.

KW Neuropeptide.

SQ SEQUENCE 5 AA; 649 MW; 71B7673B44600000 CRC64;

Query Match

21.2%; Score 7; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 Y 2

Db 2 Y 2

Search completed: April 22, 2003, 13:40:08

Job time: 11 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:37:57 ; Search time 82 Seconds  
(without alignments)  
12.564 Million cell updates/sec

Title: US-09-674-716B-9  
Perfect score: 33  
Sequence: 1 GYMS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0  
Maximum DB seq length: 5

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacterioph.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	18.2	5	10 Q99007	Q99007 hordeum vul
2	5	15.2	5	2 P83073	P83073 bacillus ce
3	3	9.1	5	13 P83308	P83308 gallus gall
4	2	6.1	4	11 Q08433	Q08433 rattus norv

## ALIGNMENTS

RESULT 1  
Q99007  
ID Q99007 PRELIMINARY; PRT; 5 AA.  
AC Q99007  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE Alpha-amylase (EC 3.2.1.1) (Fragment).  
GN AMYL.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
RN [1]\_TaxID=4513;  
RP SEQUENCE FROM N.A.  
RC STRAIN=HIMALAYA; TISSUE=ALEURONE LAYER;  
RX MEDLINE=91329704; PubMed=1831055;  
RA Jacobsen J.V., Close T.J.;  
RT "Control of transient expression of chimaeric genes by gibberellic acid and abscisic acid in protoplasts prepared from mature barley aleurone layers".  
RL Plant Mol. Biol. 16:713-721(1991).  
CC -|- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
CC -|- COPACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.  
CC -|- MISCELLANEOUS: THERE ARE AT LEAST 4 TYPES OF ALPHA-AMYLASE IN BARLEY.  
DR EMBL; X54643; CAA38455.1; -.  
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Seed; Germination;  
KW Calcium; Multigene family.  
FT NON\_TER 5  
SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;  
  
Query Match 18.2%; Score 6; DB 10; Length 5;  
Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 4 MS 5  
Db 1 MA 2  
  
RESULT 2  
P83073  
ID P83073 PRELIMINARY; PRT; 5 AA.  
AC P83073  
DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE 88 kDa protein (Fragment).  
OS Bacillus cereus.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1396;  
RN [1]\_TaxID=1396;  
RP SEQUENCE  
RC STRAIN=NCIMB 11796;  
RA Browne N., Dowds B.C.A.;  
RL Submitted (JUL-2001) to the SWISS-PROT data bank.  
FT NON\_TER 5  
SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;  
  
Query Match 15.2%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 6.7e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 4 M 4  
Db 1 M 1  
  
RESULT 3  
P83308  
ID P83308 PRELIMINARY; PRT; 5 AA.  
AC P83308  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE FMRFamide-like neuropeptide (LPLRF-amide).

OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;

RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=BRAIN;  
 RX PubMed=6137771;  
 RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;  
 RT "A novel active pentapeptide from chicken brain identified by  
 RT antibodies to FMRamide";  
 RL Nature 305:328-330(1983).  
 CC -!- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRPAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide.  
 SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;

Query Match 9.1%; Score 3; DB 13; Length 5;  
 Best Local Similarity 0.0%; Pred. NO. 6.7e+05;  
 Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 Y 2  
 Db 5 F 5

## RESULT 4

Q08433  
 ID Q08433 PRELIMINARY; PRT; 4 AA.  
 AC Q08433;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
 DE UDP-glucuronosyltransferase, microsomal (EC 2.4.1.17) (UDPGT)  
 DE (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GUNN;  
 RX MEDLINE=91282758; PubMed=1840486;  
 RA Sato H., Aono S., Kashiwamata S., Koiwai O.;  
 RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the  
 RT hyperbilirubinemic Gunn rat.";  
 RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).  
 CC -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND  
 CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND  
 CC ENDOGENOUS COMPOUNDS.  
 CC -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR  
 CC BETA-D-GLUCURONOSIDE.  
 CC -!- SUBCELLULAR LOCATION: MICROSOME.  
 DR EMBL; S38636; AAB19259.1; -;  
 KW Transferase; Glycosyltransferase; Microsome; Multigene family.  
 FT NON\_TER 1 1  
 FT NON\_TER 4 4  
 SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 6.1%; Score 2; DB 11; Length 4;  
 Best Local Similarity 0.0%; Pred. NO. 6.7e+05;  
 Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 M 4  
 Db 3 L 3

Search completed: April 22, 2003, 13:41:37  
 Job time : 82 secs



GenCore version 5.1.4.p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:32:01 ; Search time 33 Seconds  
(without alignments)  
20.189 Million cell updates/sec

Title: US-09-674-716B-9

Perfect score: 33

Sequence: 1 GYWS 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 24944

Minimum DB seq length: 0

Maximum DB seq length: 5

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq 101002.\*

1: /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID82/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	5	21	Light chain CDR H1
2	29	87.9	5	15	Murine anti-human
3	27	81.8	5	17	Humanised monoclon
4	27	81.8	5	20	Tumour antigen ant
5	27	81.8	5	22	Anti-TANGO 268 scf
6	27	81.8	5	23	Human PAPalpha spe
7	27	81.8	5	23	Anti-(MCP)-1 antib
8	25	75.8	5	16	MAB 55.1 heavy cha
9	24	72.7	4	23	Murine MC-1 antibo
10	24	72.7	5	16	ScFv(FWP51) CDR1H.

11	24	72.7	5	18	AAW23429	CDR-1 of rW12 heavy
12	24	72.7	5	20	AAW93474	mak TPC8 hypervari
13	24	72.7	5	22	AAW62861	Anti-SAP-1 antibody
14	24	72.7	5	22	ABB07357	22A5 IGM heavy cha
15	23	69.7	5	16	AAW74929	H-CDR-1 of anti-id
16	23	69.7	5	16	AAW74931	H-CDR-1 of anti-id
17	23	69.7	5	18	AAW24546	CDR #1 of r101-2 h
18	23	69.7	5	19	AAW83024	Anti-Fas MAB HFE7A
19	23	69.7	5	19	AAW62192	Mouse anti-HM1.24
20	23	69.7	5	19	AAW57588	Chimeric H chain V
21	23	69.7	5	19	AAW44181	Monoclonal antibody
22	23	69.7	5	20	AAV02550	Artificial CDR(1)
23	23	69.7	5	20	AAV05045	Tumour antigen ant
24	23	69.7	5	20	AAW83632	Mouse humanised an
25	23	69.7	5	21	AAAB19754	Erythropoietin rec
26	23	69.7	5	21	AAAB14741	Mouse anti-Fas ant
27	23	69.7	5	21	AAAB12174	Human CDR1 for IL-
28	23	69.7	5	21	AAW90891	Murine anti-Fas an
29	23	69.7	5	21	AAW92158	Murine 15B8 heavy
30	23	69.7	5	21	AAW80139	Chimeric anti-CD25
31	23	69.7	5	21	AAV77518	Antibody L chain V
32	23	69.7	5	22	AAO14422	CD25 binding prote
33	23	69.7	5	22	AAW67116	Amino acid sequenc
34	23	69.7	5	22	AAW63398	Amino acid sequenc
35	23	69.7	5	22	AAW64780	Anti-PTHrP Ab VH C
36	23	69.7	5	22	AAW76894	Human PTHrP mouse
37	23	69.7	5	22	AAW76913	Human PTHrP mouse
38	23	69.7	5	22	AAW76932	Human PTHrP mouse
39	23	69.7	5	22	AAW61281	Anti-TANGO 268 scf
40	23	69.7	5	23	ABB95192	Human joint diseas
41	23	69.7	5	23	ABB74860	Humanised anti-Fas
42	23	69.7	5	23	ABB74906	Humanised anti-Fas
43	23	69.7	5	23	AAU11183	Mouse antibody Act
44	22	66.7	5	22	AAU05472	Synthetic pentapep
45	21	63.6	5	18	AAW19405	Tissue factor pept

## ALIGNMENTS

RESULT 1  
AAV32257  
ID AAV32257 standard; Peptide; 5 AA.  
AC AAV32257;  
DT 15-FEB-2000 (first entry)

Light chain CDR H1 of mouse anti-CD23 MAb C11.

CD23; FCERII; IgE receptor; monoclonal antibody; C11; mouse;  
monoclonal antibody; chimeric antibody; humanised antibody;  
complementarity determining region; CDR; autoimmune disease;  
inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
urticaria; nephrotic syndrome; glomerulonephritis;  
inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
graft-versus-host disease; COPD; bronchitis; diabetes;  
B-cell malignancy; therapy.

Mus musculus.

WO9958679-A1.

18-NOV-1999.

07-MAY-1999; 99WO-GB01434.

09-MAY-1998; 98GB-0009839.

(GLAX ) GLAXO GROUP LTD.

PI Bonnefoy JMP, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 XX WPI; 2000-053101/04.  
 DR N-PSDB; AA234742.  
 XX  
 PT Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 XX  
 PS Claim 1; Page 40; 81pp; English.  
 XX  
 CC This sequence represents complementarity determining region 1  
 CC (CDR H1) of the heavy chain of murine anti-CD23 (FCERII) monoclonal  
 CC antibody C11 (see also AAY32263). The invention provides altered  
 CC antibodies, such as chimeric or humanised antibodies, which comprise  
 CC sufficient of the amino acid sequences of C11 light and heavy chain  
 CC CDRs (see AAY32254-59) to render them capable of binding to the CD23  
 CC type II molecule expressed on haematopoietic cells. The antibodies  
 CC are used to block soluble CD23 formation for treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (clabmed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX  
 SQ Sequence 5 AA;  
 Query Match 100.0%; Score 33; DB 21; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWMS 5  
 | | | | |  
 DB 1 GYWMS 5  
 | | | | |  
 RESULT 2  
 AAR62882  
 ID AAR62882 standard; peptide; 5 AA.  
 XX  
 AC AAR62882;  
 XX  
 DT 18-JUL-1995 (first entry)  
 XX  
 DE Murine anti-human atherosclerotic plaque MAB Z2D3 VH CDR1.  
 XX  
 KW Atherosclerosis; atherosclerotic plaque; anti-hydroxysteroid antibody;  
 KW murine monoclonal antibody; heavy chain variable region: CDR1;  
 KW complementarity determining region; imaging; plaque ablation.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..5  
 FT /label= CDR1  
 XX  
 PN WO9425053-A.  
 XX  
 PD 10-NOV-1994.  
 XX  
 PF 26-APR-1994; 94WO-US04641.  
 XX  
 PR 26-APR-1993; 93US-0053451.  
 XX  
 PA (SCOT-) SCOTGEN BIOPHARMACEUTICALS INC.  
 XX  
 PI Calenoff E, Chen FW, Dittlow CC;  
 XX  
 XX WPI; 1994-357904/44.

DR N-PSDB; AAQ78734.  
 XX  
 PT New antigen comprising hydroxy-steroid and quat. ammonium salt -  
 PT and related antibodies, useful e.g. for imaging, ablating or  
 PT treating atherosclerotic plaque, and detecting plaque specific  
 PT antibodies  
 XX  
 PS Claim 199; Page 147; 288pp; English.  
 XX  
 CC Mice were immunised with an extract of human atherosclerotic plaque,  
 CC then spleen cells were fused with SP2/01-Ag-14 myeloma cells.  
 CC Hybridomas were screened by ELISA for reactivity with the immunogen  
 CC and clone Z2D3 was isolated. The Z2D3 antibody reacts specifically  
 CC with atherosclerotic tissue; it recognises a non-protein antigen  
 CC containing cholesterol (or similar steroid that is a substrate for  
 CC cholesterol oxidase) and a quaternary ammonium salt (pref. a  
 CC phosphatidylcholine or related compound that is a substrate for  
 CC phospholipase C). The CDR sequences for the heavy and light chains  
 CC of Z2D3 were determined; peptides comprising the CDRs are claimed,  
 CC including chimeric (CDR-grafted) murine-human antibodies.  
 XX  
 SQ Sequence 5 AA;  
 Query Match 87.9%; Score 29; DB 15; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWMS 5  
 | | | | |  
 DB 1 GFWMS 5  
 | | | | |  
 RESULT 3  
 AAR97316  
 ID AAR97316 standard; peptide; 5 AA.  
 XX  
 AC AAR97316;  
 XX  
 DT 15-OCT-1996 (first entry)  
 XX  
 DE Humanised monoclonal antibody heavy chain CDR.  
 XX  
 KW Monoclonal antibody; humanised; mouse; framework region; FR; CDR;  
 KW complementary determining region; anti-carcinoembryonic antigen;  
 KW CEA; diagnosis; imaging; therapy; immune response.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO9611013-A1.  
 XX  
 PD 18-APR-1996.  
 XX  
 PF 28-SEP-1995; 95WO-US11964.  
 XX  
 PR 05-OCT-1994; 94US-0318157.  
 XX  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX  
 FT Armour KL, Hansen HJ;  
 XX  
 DR WPI; 1996-209653/21.  
 XX  
 CC New humanised anti-CEA monoclonal antibody - having engrafted murine  
 CC CDRs, used for diagnosis, imaging and therapy of CEA-producing  
 CC cancers  
 XX  
 PS Claim 5; Page 38; 62pp; English.  
 XX  
 CC New humanised monoclonal antibodies (MABs) comprising the  
 CC complementary determining regions (CDRs) of a parental murine class  
 CC III anti-carcinoembryonic (CEA) MAB engrafted to the framework  
 CC regions (FRs) of a heterologous antibody which can be derived from  
 CC any species including human, retain the anti-CEA binding specificity

CC of the parental murine Mab but are less immunogenic in a human  
 CC subject than the parental Mab. The humanised antibodies can be used  
 CC in diagnosis, imaging and therapy of CEA-producing cancers and  
 CC patients receiving the humanised antibodies and conjugates show  
 CC improved therapeutic results, decreased immune responses and  
 CC decreased immune-mediated adverse effects compared to the parent  
 CC antibody. This sequence is a murine derived CDR region which  
 CC corresponds to CDR 1 in the heavy chain of the humanised Mab. See  
 CC AAR97313-97333.

SQ Sequence 5 AA;  
 Query Match 81.8%; Score 27; DB 17; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
 | | | |  
 Db 2 YWMS 5

RESULT 4  
 AAY05038  
 ID AAY05038 standard; peptide; 5 AA.  
 AC AAY05038;  
 XX  
 DT 16-JUN-1999 (first entry)  
 DE Tumour antigen antibody heavy chain CDR1 clone F15.  
 KW Tumour antigen; antibody; CDR; complementarity determining region;  
 KW binding molecule identification; tumour-specific binding polypeptide;  
 KW cancer therapy; heavy chain.  
 OS Homo sapiens.  
 XX  
 PN WO9906834-A2.  
 PD 11-FEB-1999.  
 XX  
 PF 04-AUG-1998; 98WO-US16280.  
 XX  
 PR 04-AUG-1997; 97US-0905825.  
 PR 04-AUG-1997; 97US-1112222.  
 XX  
 PA (IXSY-) IXSYS INC.  
 XX  
 PI Huse WD, Watkins JD, Wu H;  
 XX  
 DR WPI; 1999-153951/13.  
 DR N-PSDB; AAX28215.  
 XX  
 PT Identifying binding molecules for ligands, particularly tumour  
 PT antigens - by selectively immobilising a population of binding  
 PT molecules to a solid support and screening for binding to two or  
 PT more ligands  
 XX  
 PS Claim 15; Page 60; 80pp; English.  
 XX  
 CC This sequence represents a heavy chain complementarity determining  
 CC region (CDR) from a tumour antigen specific antibody.  
 CC The invention relates to a method for identifying a binding molecule  
 CC having selective affinity for a ligand comprising: (a) selectively  
 CC immobilising a diverse population of binding molecules to a solid  
 CC support; (b) simultaneously contacting the diverse population immobilised  
 CC on the solid support with 2 or more ligands; and (c) determining at least  
 CC one binding molecule which selectively binds to one or more of the  
 CC ligands. The method allows for the rapid and efficient methods for the  
 CC identification of binding molecules which exhibit selective affinity for  
 CC one or more ligands of interest. They are used particularly for  
 CC identifying tumour-specific binding polypeptides which can be used as  
 CC targeting agents for cancer therapy that minimises impact on non-tumour

CC tissues.  
 XX  
 SQ Sequence 5 AA;  
 Query Match 81.8%; Score 27; DB 20; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
 | | | |  
 Db 2 YWMS 5

RESULT 5  
 AAB61293  
 ID AAB61293 standard; Peptide; 5 AA.  
 AC AAB61293;  
 XX  
 DT 04-APR-2001 (first entry)  
 DE Anti-TANGO 268 scFv CDR, SEQ ID NO: 61.  
 XX  
 KW Human; antibody; scFv; CDR; complementarity determining region;  
 KW TANGO 268; cardiant; cerebroprotective; cytostatic; anticoagulant;  
 KW thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;  
 KW platelet membrane glycoprotein receptor; bleeding disorder;  
 KW blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;  
 KW ischaemia; cardiovascular disease; immunological disease; liver disorder;  
 KW cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200100810-A1.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 30-JUN-2000; 2000WO-US18152.  
 XX  
 PR 30-JUN-1999; 99US-0345468.  
 PR 06-DEC-1999; 99US-0454824.  
 PR 14-FEB-2000; 2000US-0503387.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Busfield SJ, Villelall J, Jandrot-Perrus M, Vainchenker W, Gill DS;  
 PI Qian MD, Kingsbury G;  
 XX  
 DR WPI; 2001-080877/09.  
 XX  
 PT New genes encoding human platelet-expressed collagen receptor,  
 PT glycoprotein VI, and its modulators, useful for preventing, treating  
 PT and diagnosing hemorrhagic disorders, thrombotic diseases and  
 PT immunological disorders -  
 XX  
 PS Claim 31; Page 102; 227pp; English.  
 XX  
 CC The present sequence is given in a specification relating to an isolated  
 CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor  
 CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides  
 CC and polypeptides and their modulators, e.g. antisense nucleic acids,  
 CC ribozymes and antibodies, are useful for preventing, treating and  
 CC diagnosing disorders associated with aberrant expression or activity of  
 CC GPVI. These disorders include bleeding disorders  
 CC (e.g. thrombocytopenia), blood vessel injury, thrombotic disorders  
 CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic  
 CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and  
 CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial  
 CC infarction), immunological diseases (e.g. platelet disorder) and  
 CC embryonic liver disorders. Preferably they are used to prevent acute  
 CC cardiac ischaemia following angioplasty and metastatic cancers,  
 CC especially of the colon and liver.  
 XX

SQ Sequence 5 AA;  
 Query Match 81.8%; Score 27; DB 22; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
 ||||  
 Db 2 YWMS 5

RESULT 6  
 AA017609  
 ID AA017609 standard; Peptide; 5 AA.  
 XX AC AA017609;  
 XX AC AA017609;  
 XX 08-AUG-2002 (first entry)  
 DE Human PAPalpha specific VL region from VH50 CDR1 peptide.  
 XX Human; PAPalpha; fibroblast activating protein alpha; antibody; Ab;  
 KW gene therapy; cancer; wound healing; inflammation; cytostatic.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200168708-A2.  
 PN  
 XX 20-SEP-2001.  
 PD  
 XX 16-MAR-2001; 2001WO-EP04716.  
 PF  
 XX 17-MAR-2000; 2000DE-1013286.  
 PR  
 XX 11-SEP-2000; 2000GB-0022216.  
 PR  
 XX (BOEH ) BOEHRINGER INGELHEIM PHARMA KG.  
 PA  
 XX Park J, Garin-Chesa P, Pfizenmaier K, Moosmayer D, Mersmann M,  
 PI Schmidt A;  
 PI  
 XX WPI; 2002-041180/05.  
 DR N-PSDB; AAL46554.  
 DR  
 XX New human humanized antibody that specifically binds to fibroblasts  
 PT activating protein alpha, useful for treating cancer or tumor, and for  
 PT imaging tumors associated with activated stromal fibroblasts, e.g. lung  
 PT or breast cancer -  
 XX  
 XX Disclosure; Fig 6C; 109pp; English. ✓  
 PS  
 XX The present invention relates to a human or humanised antibody (Ab) which  
 CC specifically binds to fibroblast activating protein alpha (PAPalpha). The  
 CC antibodies are useful for preparing a composition for the treatment of  
 CC cancer, and for imaging tumours associated with activated stromal  
 CC fibroblasts, such as colorectal cancer, non-small-cell lung cancer,  
 CC breast cancer, head and neck cancer, ovarian cancer, lung cancer, bladder  
 CC cancer, pancreatic cancer and metastatic brain cancer, and diseases  
 CC associated with the same, such as inflammation and wound healing. The  
 CC present sequence is a peptide described in the exemplification of the  
 CC invention.  
 CC  
 SQ Sequence 5 AA;  
 Query Match 81.8%; Score 27; DB 23; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
 ||||  
 Db 2 YWMS 5

RESULT 7

AAU5739  
 ID AAU5739 standard; Peptide; 5 AA.  
 XX AC AAU5739;  
 XX 08-MAY-2002 (first entry)  
 DT  
 XX Anti-(MCP)-1 antibody heavy chain protein hypervariable region CDR1.  
 DE  
 XX Human; antibody; MCP; monocyte chemoattractant; antiasthmatic;  
 KW antiallergic; antiinflammatory; idiopathic thrombocytopaenia;  
 KW immunosuppressive; cytostatic; vasotropic; antiarteriosclerotic;  
 KW antirheumatic; antiarthritic; osteopathic; antigen-binding site;  
 KW immunoglobulin heavy chain; monocyte migration; T cell migration;  
 KW CC-type chemokine; etotaxin; allergic rhinitis; cancer;  
 KW hypersensitivity response; allergic contact dermatitis;  
 KW inflammatory disease; asthma; psoriasis; COPD; osteoporosis;  
 KW inflammatory bowel disease; multiple sclerosis; autoimmune disease;  
 KW rheumatoid arthritis; diabetes; systemic lupus erythematosus;  
 KW bone disease; osteoporosis; osteoarthritis; periodontal disease;  
 KW hematological disorder; haemolytic anaemia; graft rejection;  
 KW leucocyte infiltration; restenosis; arteriosclerosis; CDR1.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200202640-A2.  
 PN  
 XX 10-JAN-2002.  
 PD  
 XX 29-JUN-2001; 2001WO-EP07468.  
 PF  
 XX 30-JUN-2000; 2000GB-0016138.  
 PR  
 XX (NOVS ) NOVARTIS AG.  
 PA  
 XX (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 XX  
 XX Hiestand P, Hofstetter H, Payne TG, Urfer R, Di Padova FE;  
 PI WPI; 2002-164525/21.  
 DR  
 XX New compound that binds human monocyte chemoattractant protein-1,  
 PT useful e.g. for treating inflammation, comprises immunoglobulin chains  
 PT with specific hypervariable regions -  
 PT  
 XX Claim 1; Page 35; 42pp; English.  
 PS  
 XX This invention relates to a molecule that binds MCP-1 (human monocyte  
 CC chemoattractant protein-1). The molecule has at least one antigen  
 CC binding site comprising at least one immunoglobulin (Ig) heavy and/or  
 CC light chain variable domain (Vh, Vl). The antibody of the invention  
 CC inhibits binding of MCP-1 (and also etotaxin) to specific receptors, so  
 CC preventing migration of monocytes and T cells. These are human  
 CC antibodies with very high affinity for MCP-1, but no significant  
 CC reaction with other human CC-type chemokines or MCP-1 from other  
 CC species. The MCP-1 binding molecules, specifically antibodies that cross  
 CC -react with etotaxin, are used to prevent or treat MCP-1 or etotaxin-  
 CC mediated disorders, particularly those that involve migration or  
 CC activation of monocytes and T cells, e.g. allergies (allergic rhinitis,  
 CC hypersensitivity responses, allergic contact dermatitis), inflammatory  
 CC diseases (asthma, psoriasis, COPD, inflammatory bowel disease,  
 CC multiple sclerosis), autoimmune disease (rheumatoid arthritis, diabetes,  
 CC systemic lupus erythematosus), diseases of bone and cartilage  
 CC (osteoporosis, osteoarthritis, periodontal disease) haematological  
 CC disorders (haemolytic anaemia, idiopathic thrombocytopaenia) graft  
 CC rejection, cancers that include leucocyte infiltration, (re)stenosis,  
 CC arteriosclerosis, osteoporosis and many other diseases listed in the  
 CC specification. The present sequence represents the human anti-(MCP)  
 CC antibody heavy chain sequence hypervariable domain peptide CDR1 of the  
 CC invention.  
 CC  
 SQ Sequence 5 AA;  
 Query Match 81.8%; Score 27; DB 23; Length 5;

```

Best Local Similarity 100.0%; Pred. No. 7.8e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

QY 2 YWMS 5
DB 2 YWMS 5

RESULT 8
AAR76079
ID AAR76079 standard; Peptide; 5 AA.
XX
AC AAR76079;
XX
DT 21-NOV-1995 (first entry)
XX
DE MAb 55.1 heavy chain CDR1.
XX
XX Antigen binding structure; complementarity determining region; CDR;
KW CA55.1; colorectal cancer; tumor-associated antigen; hybridoma;
KW monoclonal antibody; MAb; immunotherapy; therapy; diagnosis;
KW transgenic animal; transgenic plant; antibody engineering;
KW humanized antibody; immunotoxin.
XX
OS Mus sp.
XX
XX WO9515382-A.
XX
XX 08-JUN-1995.
PD
XX 29-NOV-1994; 94WO-GB02610.
XX
PR 03-JUN-1994; 94GB-0011089.
PR 03-DEC-1993; 93GB-0024819.
XX
PA (ZENE ) ZENECA LTD.
XX
PI Blakey DC, Boot C, Copley CG, Hall SM, Paterson DS;
PI Rose MS, Wright AF;
XX
XX WPI; 1995-215262/28.
DR
XX Antigen binding structures containing CDRs recognising the CA55.1
PT antigen - produced by hybridomas and host cells, for use in the
PT diagnosis and therapy of cancer
XX
PS Claim 2; Page 96; 121pp; English.
XX
CC An antigen binding structure is based on the CDRs (given in AAR76078-
CC 84) of the heavy and light chains of MAb 55.1 (ECACC 93081901),
CC which recognises the colorectal tumor-associated antigen CA55.1.
CC It is optionally humanized and in the form F(ab')2, F(ab)'', Fab,
CC Fv, scFv or V-min, and is produced in transgenic animals or plants.
XX
SQ Sequence 5 AA;

Query Match 75.8%; Score 25; DB 16; Length 5;
Best Local Similarity 75.0%; Pred. No. 7.8e+05; Indels 0; Gaps 0;
Matches 3; Conservative 1; Mismatches 0;

QY 1 GYWM 4
DB 1 GYWI 4

RESULT 9
AAE22200
ID AAE22200 standard; peptide; 4 AA.
XX
AC AAE22200;
XX
XX 25-JUL-2002 (first entry)
DT
XX

```

```

DE Murine MC-1 antibody heavy chain variable domain (VH) CDR1 peptide.
XX
XX Chemokine construct; human immunodeficiency virus 1; allergic disease;
KW skin disease; immunological disorder; autoimmune disease; psoriasis;
KW multiple sclerosis; rheumatoid arthritis; inflammatory bowel disease;
KW diabetes; skin inflammation; atopic dermatitis; inflammatory disease;
KW inflammatory renal disease; HIV-1; transplant rejection; murine; CDR1;
KW heavy chain variable domain; VH; complementarity determining region 1;
KW MC-1; antibody.
XX
XX Mus sp.
XX
XX WO200220615-A2.
XX
XX 14-MAR-2002.
PD
XX 10-SEP-2001; 2001WO-EP10433.
PP
XX 08-SEP-2000; 2000EP-0119694.
PR
XX 05-SEP-2001; 2001US-0948004.
XX
PA (MICR-) MICROMET AG.
XX
XX Mack M, Schloendorff D, Spring M;
PI
XX WPI; 2002-362240/39.
DR
XX Use of an antibody and/or chemokine construct that binds to a chemokine
PT receptor, for eliminating cells latently infected with primate
PT immunodeficiency virus, or treating, preventing and alleviating immune
PT disorders -
XX
XX Disclosure; Page 116; 117pp; English.
XX
CC The invention relates to the use of an antibody and/or chemokine
CC construct that binds a chemokine receptor for preparing a pharmaceutical
CC composition for eliminating cells latently infected with a primate
CC immunodeficiency virus (e.g. human immunodeficiency virus (HIV-1)). They
CC are used in gene therapy and as vaccines. The antibody and/or chemokine
CC construct is also used for preparing a pharmaceutical composition for
CC treating, preventing and/or alleviating immunological disorders
CC including autoimmune diseases (e.g. multiple sclerosis, type I diabetes
CC and rheumatoid arthritis), allergic diseases, skin diseases (e.g. skin
CC inflammation, atopic dermatitis and psoriasis), inflammatory diseases
CC such as inflammatory joint diseases (chronic arthritis), inflammatory
CC renal diseases and inflammatory bowel diseases and graft versus host
CC and transplant rejections. The present sequence is murine MC-1 antibody
CC heavy chain variable domain (VH) complementarity determining region 1
CC (CDR1) peptide.
XX
XX Sequence 4 AA;
SQ

Query Match 72.7%; Score 24; DB 23; Length 4;
Best Local Similarity 75.0%; Pred. No. 7.8e+05; Indels 0; Gaps 0;
Matches 3; Conservative 1; Mismatches 0;

QY 2 YWMS 5
DB 1 YWMN 4

RESULT 10
AAR85499
ID AAR85499 standard; Peptide; 5 AA.
XX
XX AAR85499;
AC
XX 16-MAR-1996 (first entry)
DT
XX SCFv(FWP51) CDR1H.
DE
XX Single chain antibody; scFv; antibody engineering; antitumour;
KW tumour antigen binding; cytotoxic T-lymphocyte; cell targeting;

```

KW monoclonal antibody; cancer; adoptive immunotherapy;  
 XX complementarity determining region; CDR.  
 OS Synthetic.  
 XX WO9530014-A1.  
 XX 09-NOV-1995.  
 XX 20-APR-1995; 95WO-EP01494.  
 XX 02-MAY-1994; 94EP-0810244.  
 XX (CIBA ) CIBA GEIGY AG.  
 XX Groner B, Moritz D;  
 XX WPI; 1995-393085/50.  
 XX New bifunctional proteins for use in killing tumour cells - contg. a  
 PT tumour antigen binding domain, a hinge region and a zeta chain  
 PT derived from a T-cell antigen receptor  
 XX Disclosure; Page 6; 46pp; English.  
 XX Complementarity determining regions (CDRs) CDR1H, CDR2H and CDR3H  
 CC of the heavy chain variable region and CDR3L (AAR85502-04) are components  
 CC respectively, of single chain antibody scFv(FWP51) (AAR85495). The CDRs  
 CC are derived from mouse monoclonal antibody FWP51 (ECACC 90112118)  
 CC directed to the human growth factor receptor HER2. The scFv is  
 CC used as the tumour antigen binding domain of a bifunctional protein  
 CC that is expressed in cytotoxic T-lymphocytes (CTLs), allowing  
 CC adoptive immunotherapy of cancer.  
 XX Sequence 5 AA;  
 SQ  
 Query Match 72.7%; Score 24; DB 16; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 YWMS 5  
 DB |||:  
 2 YWMN 5  
 RESULT 11  
 ID AAW23429 standard; peptide; 5 AA.  
 XX AAW23429;  
 XX 23-APR-1998 (first entry)  
 XX CDR-1 of rW12 heavy chain.  
 XX Antibody; complementarity determining region; CDR; heavy chain; rat; CEA;  
 KW rW12; anti-carcinoembryonic antigen; vaccine; immune response; cancer;  
 KW therapy; idiotype region.  
 XX Rattus sp.  
 XX WO9734636-A1.  
 XX 25-SEP-1997.  
 XX 19-MAR-1997; 97WO-US04696.  
 XX 20-MAR-1996; 96US-0013708.  
 XX (IMMU- ) IMMUNOMEDICS INC.  
 XX Hansen H, Leung S, Losman MJ;  
 PI  
 monoclinal antibody; cancer; adoptive immunotherapy;  
 XX complementarity determining region; CDR.  
 OS Synthetic.  
 XX WO9530014-A1.  
 XX 09-NOV-1995.  
 XX 20-APR-1995; 95WO-EP01494.  
 XX 02-MAY-1994; 94EP-0810244.  
 XX (CIBA ) CIBA GEIGY AG.  
 XX Groner B, Moritz D;  
 XX WPI; 1995-393085/50.  
 XX New bifunctional proteins for use in killing tumour cells - contg. a  
 PT tumour antigen binding domain, a hinge region and a zeta chain  
 PT derived from a T-cell antigen receptor  
 XX Disclosure; Page 6; 46pp; English.  
 XX Complementarity determining regions (CDRs) CDR1H, CDR2H and CDR3H  
 CC of the heavy chain variable region and CDR3L (AAR85502-04) are components  
 CC respectively, of single chain antibody scFv(FWP51) (AAR85495). The CDRs  
 CC are derived from mouse monoclonal antibody FWP51 (ECACC 90112118)  
 CC directed to the human growth factor receptor HER2. The scFv is  
 CC used as the tumour antigen binding domain of a bifunctional protein  
 CC that is expressed in cytotoxic T-lymphocytes (CTLs), allowing  
 CC adoptive immunotherapy of cancer.  
 XX Sequence 5 AA;  
 SQ  
 Query Match 72.7%; Score 24; DB 16; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 YWMS 5  
 DB |||:  
 2 YWMN 5  
 RESULT 12  
 ID AAW93474 standard; Protein; 5 AA.  
 XX AAW93474;  
 XX 11-JUN-1999 (first entry)  
 XX mAk TTC8 hypervariable domain heavy chain CDR1 protein fragment.  
 XX Monoclonal antibody; hypervariable domain; heavy chain; CDR1; neutralise;  
 KW Clostridium difficile; epitope; ligand domain; immunotherapy; disease;  
 KW translocation domain; catalytic domain; humanised antibody; enterotoxin;  
 KW toxin A; toxin B; pseudomembranous colitis.  
 XX Synthetic.  
 XX DE19739685-A1.  
 XX 11-MAR-1999.  
 XX 10-SEP-1997; 97DE-1039685.  
 XX 10-SEP-1997; 97DE-1039685.  
 XX (VEIC/) VON EICHEL-STREIBER C.  
 XX Moos M, Von Eichel-Streiber C;  
 XX WPI; 1999-182094/16.  
 XX N-PSDB; AAX23235.  
 XX Monoclonal antibodies specific for Clostridium difficile toxins -  
 PT especially humanised antibodies for treating pseudomembranous  
 PT colitis  
 XX Claim 20; Page 8; 14pp; German.

CC This invention describes a novel monoclonal antibody that is directed  
 CC against a Clostridium difficile toxin and recognises and neutralises  
 CC an epitope in the ligand domain, translocation domain or catalytic  
 CC domain of the toxin. Humanised antibodies are described which are  
 CC expressed in plants and can be used for immunotherapy of diseases caused  
 CC by Clostridium difficile enterotoxin (toxin A) or cytotoxin (toxin B),  
 CC especially pseudomembranous colitis.

XX SQ Sequence 5 AA;  
 Query Match 72.7%; Score 24; DB 20; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YWMS 5  
 |||:  
 Db 2 YWMN 5

RESULT 13  
 AAB62861  
 ID AAB62861 standard; peptide; 5 AA.

XX AC AAB62861;  
 XX 15-MAY-2001 (first entry)

XX DE Anti-SAF-1 antibody 13G3 H chain V region CDR peptide SEQ ID 3.

XX KW Sialoadhesin factor-1; SAF-1; anti-SAF-1 monoclonal antibody; 13G3;  
 KW 11G4; osteomyelitis; wound infection; sepsis; listeria monocytogene;  
 KW Salmonella; Leishmania; heavy chain variable region; mouse;  
 KW complementarity determining region; CDR.

XX OS Mus musculus.

XX XN WO200112646-A1.

XX XX 22-FEB-2001.

XX XX 18-AUG-2000; 2000WO-US22663.

XX XX 19-AUG-1999; 99US-0149753.

XX PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX PI Erickson-Miller CL, Holmes SD, King AG, Taylor AH;

XX DR WPI; 2001-211197/21.

XX DR N-PSDB; AAF62142.

XX PT New sialoadhesin factor-1 agonist antibody having the characteristic of  
 PT monoclonal antibody 13G3 or 11G4 for treating osteomyelitis, deep  
 PT seated wound infections and antibiotic resistant bacterial infections

XX PS Claim 22; Fig 1; 43pp; English.

XX CC This invention relates to sialoadhesin factor-1 (SAF-1) monoclonal  
 CC antibodies 13G3 and 11G4. A neurophil-priming dose of SAF-1 agonist  
 CC antibody or a pharmaceutical composition comprising the antibody is  
 CC useful for treating or preventing an infectious disease state such as  
 CC osteomyelitis, deep seated wound infections, sepsis, antibiotic resistant  
 CC bacterial infection, viral infections and those involving intracellular  
 CC pathogens and parasites such as listeria monocytogenes, Salmonella and  
 CC Leishmania in a mammal. Nucleic acid sequences encoding the variable  
 CC light chain and heavy chain peptide sequences are useful for mutagenic  
 CC introduction of specific changes within the nucleic acid sequences  
 CC encoding the (complementarity determining regions) CDRs or framework  
 CC regions, and for incorporation of the resulting modified or fusion  
 CC nucleic acid sequence into a plasmid for expression. The CDR-encoding  
 CC regions comprising silent mutations are used in construction of humanized

CC antibodies or other engineered antibodies. The present sequence  
 CC represents anti-SAF-1 monoclonal antibody 13G3 heavy chain variable  
 CC region complementarity determining region peptide sequence.

XX SQ Sequence 5 AA;

Query Match 72.7%; Score 24; DB 22; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YWMS 5  
 |||:  
 Db 2 YWMN 5

RESULT 14  
 ABB07357  
 ID ABB07357 standard; peptide; 5 AA.

XX AC ABB07357;

XX DT 09-APR-2002 (first entry)

XX DE 22A5 IGM heavy chain variable region CDR1 sequence.

XX KW Immunoglobulin; Ig; reperfusion; ischaemic injury; pathogenic; CDR;  
 XX vasotrophic; antibody; 22A5; IGM; complementarity determining region.

XX OS Mus musculus.

XX PN WO2001193892-A1.

XX PD 13-DEC-2001.

XX PF 08-JUN-2001; 2001WO-US18510.

XX PR 08-JUN-2000; 2000US-210272P.

XX PA (BLOO-) CENT BLOOD RES INC.

XX PI Carroll MC, Moore FD, Hechtman HB;

XX XX WPI; 2002-139643/18.

XX DR N-PSDB; ABA94514.

XX PT Treating or preventing immunoglobulin-mediated reperfusion or ischemic  
 PT injury resulting from e.g. angioplasty or bypass surgery, by  
 PT administering an inhibitor of pathogenic immunoglobulin and ischemic  
 PT antigen interaction

XX PS Claim 54; Page 72; 77pp; English.

XX CC The invention provides a method of treating or preventing immunoglobulin-  
 CC mediated reperfusion or ischemic injury in a subject. The method involves  
 CC administering an inhibitor of an interaction between a pathogenic  
 CC immunoglobulin, and an ischemic-specific antigen or a component of the  
 CC complement pathway. The method is useful for treating or preventing  
 CC immunoglobulin-mediated reperfusion or ischemic injury, in particular,  
 CC injury that results following a naturally occurring episode, a surgical  
 CC procedure (e.g. angioplasty, stenting procedure, atherectomy or bypass  
 CC surgery), or in an injury that occurs in a cardiovascular tissue. The  
 CC method is also useful for treating or preventing, in a subject, tissue  
 CC damage following reperfusion caused by a pathogenic immunoglobulin, e.g.  
 CC pathogenic IGMs. The present sequence represents the complementarity  
 CC determining region (CDR) of a heavy chain variable region of 22A5 IGM, a  
 CC pathogenic immunoglobulin.

XX SQ Sequence 5 AA;

Query Match 72.7%; Score 24; DB 23; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YWMS 5  
Db 2 YWMN 5

RESULT 15

AAR74929  
ID AAR74929 standard; peptide; 5 AA.

XX AC AAR74929;  
XX DT 19-JAN-1996 (first entry)

XX H-CDR-1 of anti-idiotypic antibody against human anticancer antibody.  
XX DE Antibody; cancer; CDR; heavy chain; light chain; immunoglobulin;  
XX KW complementarity determining region.

XX OS Mus sp.

XX JP07101999-A.

XX 18-APR-1995.

XX 06-OCT-1993; 93JP-0272950.

XX 06-OCT-1993; 93JP-0272950.

XX (HAGI/) HAGIWARA Y.

XX WPI; 1995-182987/24.

XX Novel anti-idiotypic antibody against an human anticancer monoclonal

PT antibody - and DNA sequences encoding the antibody, useful in

PT pharmacology, medicine and biochemical fields.

XX Claim 1; Page 2; 28pp; Japanese.

XX A new anti-idiotypic antibody against a human anticancer monoclonal  
CC antibody is claimed. This antibody contains in its heavy chain 3  
CC complementarity determining regions CDR1 (AAR74929-R74931), CDR2  
CC (AAR74932-R74935) and CDR3 (AAR74936-R74939), this is also true of the  
CC light chain which has its own CDR1 (AAR74944-R74946 and AAR85774), CDR2  
CC (AAR74947-R74949) and CDR3 (AAR74950-R74954). The antibody and DNA  
CC encoding it are useful in pharmacological, medical and biochemical  
CC fields.

SQ Sequence 5 AA;

Query Match 69.7%; Score 23; DB 16; Length 5;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YWM 4  
Db 2 YWM 4

Search completed: April 22, 2003, 13:39:53  
Job time : 35 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:41:42 ; Search time 15 Seconds  
(without alignments)  
26.710 Million cell updates/sec

Title: US-09-674-716B-9

Perfect score: 33

Sequence: 1 GYMMS 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 4813

Minimum DB seq length: 0

Maximum DB seq length: 5

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US05\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	81.8	5	9	US-09-977-797A-76
2	27	81.8	5	10	US-09-832-312-61
3	27	81.8	5	10	US-09-253-794-23
4	24	72.7	5	9	US-09-968-561A-265
5	24	72.7	5	9	US-09-155-106-1
6	24	72.7	5	10	US-09-192-854-151
7	23	69.7	5	9	US-09-423-800-62
8	23	69.7	5	9	US-09-269-921-6
9	23	69.7	5	9	US-09-977-797A-90
10	23	69.7	5	9	US-10-182-018-62
11	23	69.7	5	10	US-09-748-960-12
12	23	69.7	5	10	US-09-832-312-49
13	23	69.7	5	10	US-09-770-002-1
14	20	60.6	5	9	US-09-291-417-40
15	19	57.6	5	9	US-10-029-301-4
16	18	54.5	4	9	US-09-750-754-14
17	18	54.5	4	12	US-10-156-820-31
18	18	54.5	5	9	US-09-750-754-3
19	18	54.5	5	9	US-09-750-754-15

20	18	54.5	5	12	US-10-156-820-34	Sequence 34, Appl
21	17	51.5	5	9	US-09-977-797A-88	Sequence 88, Appl
22	16	48.5	4	8	US-08-484-409-5	Sequence 5, Appl
23	16	48.5	4	9	US-09-882-781-15	Sequence 15, Appl
24	16	48.5	4	10	US-09-982-172-194	Sequence 194, Appl
25	16	48.5	4	10	US-09-929-924-41	Sequence 41, Appl
26	16	48.5	5	9	US-09-963-298-8	Sequence 8, Appl
27	16	48.5	5	9	US-09-291-417-38	Sequence 38, Appl
28	15	45.5	4	9	US-10-165-015-10	Sequence 10, Appl
29	15	45.5	4	10	US-09-853-918-19	Sequence 19, Appl
30	15	45.5	4	10	US-09-977-831-10	Sequence 10, Appl
31	15	45.5	5	9	US-10-105-930-67	Sequence 67, Appl
32	15	45.5	5	9	US-10-171-077-2	Sequence 2, Appl
33	15	45.5	5	9	US-09-788-006-95	Sequence 95, Appl
34	15	45.5	5	9	US-09-788-006-96	Sequence 96, Appl
35	15	45.5	5	9	US-09-788-006-97	Sequence 97, Appl
36	15	45.5	5	9	US-09-977-797A-80	Sequence 80, Appl
37	15	45.5	5	10	US-09-952-768-69	Sequence 69, Appl
38	15	45.5	5	10	US-09-952-768-70	Sequence 70, Appl
39	15	45.5	5	10	US-09-730-857-84	Sequence 84, Appl
40	14	42.4	4	9	US-09-264-516A-29	Sequence 29, Appl
41	14	42.4	4	9	US-09-758-426-58	Sequence 58, Appl
42	14	42.4	4	9	US-09-769-145-74	Sequence 74, Appl
43	14	42.4	4	9	US-09-758-198-58	Sequence 58, Appl
44	14	42.4	4	9	US-09-185-908-24	Sequence 24, Appl
45	14	42.4	4	9	US-10-119-537-8	Sequence 8, Appl

## ALIGNMENTS

### RESULT 1

US-09-977-797A-76  
; Sequence 76, Application US/09977797A  
; Publication No. US20030044772A1  
; GENERAL INFORMATION:  
; APPLICANT: Watkins, Jeffrey D.  
; APPLICANT: Huse, William D.  
; APPLICANT: Wu, Herren  
; TITLE OF INVENTION: Methods for Identifying Ligand Specific Binding Molecules  
; FILE REFERENCE: AME-06805  
; CURRENT APPLICATION NUMBER: US/09/977,797A  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 09/129,026  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 08/905,825  
; PRIOR FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 136  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 76  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-797A-76

Query Match 81.8%; Score 27; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred.No.2.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy. 2 YWMS 5

Db 2 YWMS 5

### RESULT 2

US-09-832-312-61  
; Sequence 61, Application US/09832312  
; Patent No. US20010049829A1  
; GENERAL INFORMATION:  
; APPLICANT: Busfield et al.  
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF  
; FILE REFERENCE: 7853-234  
; CURRENT APPLICATION NUMBER: US/09/832,312

;; CURRENT FILING DATE: 2001-04-09  
;; PRIOR APPLICATION NUMBER: 09/610,118  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: 09/503,387  
;; PRIOR FILING DATE: 2000-02-14  
;; PRIOR APPLICATION NUMBER: 09/454,824  
;; PRIOR FILING DATE: 1999-12-06  
;; PRIOR APPLICATION NUMBER: 09/345,468  
;; PRIOR FILING DATE: 1999-06-30  
;; NUMBER OF SEQ ID NOS: 78  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 61  
;; LENGTH: 5  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-832-312-61

Query Match 81.8%; Score 27; DB 10; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
|||  
Db 2 YWMS 5

RESULT 3  
US-09-253-794-23  
;; Sequence 23, Application US/09253794  
;; Patent No. US20020018750A1  
;; GENERAL INFORMATION:  
;; APPLICANT: HANSEN, Hans J.  
;; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
;; NUMBER OF SEQUENCES: 58  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Foley & Lardner  
;; STREET: 3000 K Street, N.W., Suite 500  
;; CITY: Washington  
;; STATE: D. C.  
;; COUNTRY: USA  
;; ZIP: 20007-5109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/253,794  
;; FILING DATE: 22-Feb-1999  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/318,157  
;; FILING DATE: 05-OCT-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: SAXE, Bernhard D.  
;; REGISTRATION NUMBER: 28,665  
;; REFERENCE/DOCKET NUMBER: 18733/464  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202)672-5300  
;; TELEFAX: (202)672-5399  
;; TELEX: 904136  
;; INFORMATION FOR SEQ ID NO: 23:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: <Unknown>  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 23:

US-09-253-794-23

Query Match 81.8%; Score 27; DB 10; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 2 YWMS 5  
|||  
Db 2 YWMS 5

RESULT 4  
US-09-968-561A-265  
;; Sequence 265, Application US/09968561A  
;; Patent No. US2002016462A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Tomlinson, Ian M  
;; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands  
;; FILE REFERENCE: 8039/1073B  
;; CURRENT APPLICATION NUMBER: US/09/968,561A  
;; CURRENT FILING DATE: 2001-10-01  
;; PRIOR APPLICATION NUMBER: GB 9722131.1  
;; PRIOR FILING DATE: 1997-10-20  
;; PRIOR APPLICATION NUMBER: US 60/065,248  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: US 60/066,729  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: PCT/GB98/03135  
;; PRIOR FILING DATE: 1998-10-20  
;; PRIOR APPLICATION NUMBER: US 09/511,939  
;; PRIOR FILING DATE: 2000-02-24  
;; NUMBER OF SEQ ID NOS: 350  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 265  
;; LENGTH: 5  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-968-561A-265

Query Match 72.7%; Score 24; DB 9; Length 5;  
Best Local Similarity 80.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYMS 5  
||:|  
Db 1 GYMS 5

RESULT 5  
US-09-155-106-1  
;; Sequence 1, Application US/09155106  
;; Publication No. US20030054003A1  
;; GENERAL INFORMATION:  
;; APPLICANT:  
;; TITLE OF INVENTION: HUMANIZATION OF AN ANTI-CARCINOEMBRYONIC  
;; TITLE OF INVENTION: ANTIGEN ANTI-ID10TYPE ANTIBODY AND USE AS A TUMOR VACCINE  
;; TITLE OF INVENTION: AND FOR TARGETING APPLICATIONS  
;; NUMBER OF SEQUENCES: 45  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/155,106  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US97/04696  
;; FILING DATE: 19-MAR-1997  
;; APPLICATION NUMBER: US 60/013,708  
;; FILING DATE: 20-MAR-1996  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids

;  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-155-106-1

Query Match 72.7%; Score 24; DB 9; Length 5;  
Best Local Similarity 75.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
|||  
DB 2 YWMT 5

## RESULT 6

US-09-192-854-151  
; Sequence 151, Application US/09192854  
; Patent No. US20020068276A1  
; GENERAL INFORMATION:  
; APPLICANT: Winter, Greg  
; APPLICANT: Tomlinson, Ian  
; TITLE OF INVENTION: Methods for Selecting Functional Peptides  
; FILE REFERENCE: 3789/72916  
; CURRENT APPLICATION NUMBER: US/09/192,854  
; CURRENT FILING DATE: 1998-11-17  
; EARLIER APPLICATION NUMBER: 60/066,729  
; EARLIER FILING DATE: 1997-11-21  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 151  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-192-854-151

Query Match 72.7%; Score 24; DB 10; Length 5;  
Best Local Similarity 80.0%; Pred. No. 2.7e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5  
|||  
DB 1 GYMS 5

## RESULT 7

US-09-423-800-62  
; Sequence 62, Application US/09423800  
; Patent No. US20020165363A1  
; GENERAL INFORMATION:  
; APPLICANT: SATO, KOH  
; APPLICANT: TSUNENARI, TOSHIKI  
; APPLICANT: ISHII, KIMIE  
; TITLE OF INVENTION: CACHEXIA REMEDY  
; FILE REFERENCE: 04853-0036  
; CURRENT APPLICATION NUMBER: US/09/423,800  
; CURRENT FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: PCT/JP98/02116  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: JP 125505/1997  
; PRIOR FILING DATE: 1997-05-15  
; PRIOR APPLICATION NUMBER: JP 194445/1997  
; PRIOR FILING DATE: 1997-07-18  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 62  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-423-800-62

Query Match 69.7%; Score 23; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 YWM 4  
|||  
DB 2 YWM 4

## RESULT 8

US-09-269-921-6  
; Sequence 6, Application US/09269921  
; Publication No. US20030045691A1  
; GENERAL INFORMATION:  
; APPLICANT: Ono, Koichiro  
; APPLICANT: Ohtomo, Toshihiko  
; APPLICANT: Tsuchiya, Masayuki  
; APPLICANT: Yoshimura, Yasuo  
; APPLICANT: Koishihara, Yasuo  
; TITLE OF INVENTION: RESHAPED HUMAN ANTI-HM 1.24 ANTIBODY  
; FILE REFERENCE: 35029-20007.00  
; CURRENT APPLICATION NUMBER: US/09/269,921  
; CURRENT FILING DATE: 1999-04-01  
; EARLIER APPLICATION NUMBER: PCT/JP97/03553  
; EARLIER FILING DATE: 1997-10-03  
; EARLIER APPLICATION NUMBER: JP 8-264756  
; EARLIER FILING DATE: 1996-10-04  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: murine  
; FEATURE:  
; OTHER INFORMATION: Amino acid sequence of CDR(1) of H chain V region of  
; OTHER INFORMATION: mouse anti-HM 1.24 antibody  
US-09-269-921-6

Query Match 69.7%; Score 23; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWM 4  
|||  
DB 2 YWM 4

## RESULT 9

US-09-977-797A-90  
; Sequence 90, Application US/09977797A  
; Publication No. US20030044772A1  
; GENERAL INFORMATION:  
; APPLICANT: Watkins, Jeffrey D.  
; APPLICANT: Huse, William D.  
; APPLICANT: Wu, Herren  
; TITLE OF INVENTION: Methods for Identifying Ligand Specific Binding Molecules  
; FILE REFERENCE: AME-06805  
; CURRENT APPLICATION NUMBER: US/09/977,797A  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 09/129,026  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 08/905,825  
; PRIOR FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 136  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 90  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-797A-90

Query Match 69.7%; Score 23; DB 9; Length 5;  
Best Local Similarity 75.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
|||  
Db 2 YWIS 5

## RESULT 10

US-10-182-018-62  
; Sequence 62, Application US/10182018  
; Publication No. US2003004911A1

; GENERAL INFORMATION:

; APPLICANT: CHUGAI SEIYAKU KABUSHIKI KAISHA

; TITLE OF INVENTION: THERAPEUTIC AND PREVENTIVE AGENTS FOR DENTAL DISEASES

; FILE REFERENCE: PH-1092-PCT

; CURRENT APPLICATION NUMBER: US/10/182,018

; CURRENT FILING DATE: 2002-07-24

; PRIOR APPLICATION NUMBER: JP 2000-83034

; PRIOR FILING DATE: 2000-01-25

; NUMBER OF SEQ ID NOS: 75

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 62

; LENGTH: 5

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-182-018-62

Query Match 69.7%; Score 23; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWM 4  
|||  
Db 2 YWM 4

## RESULT 11

US-09-748-960-12

; Sequence 12, Application US/09748960

; Patent No. US20010046496A1

; GENERAL INFORMATION:

; APPLICANT: Brettman, Lee R.

; APPLICANT: Fox, Judith A.

; APPLICANT: Allison, David Edward

; TITLE OF INVENTION: Method of Administering an Antibody

; FILE REFERENCE: 1855.2007-001

; CURRENT APPLICATION NUMBER: US/09/748,960

; CURRENT FILING DATE: 2000-12-27

; PRIOR APPLICATION NUMBER: US 09/550,082

; PRIOR FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 5

; TYPE: PRT

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (1)....(5)

; OTHER INFORMATION: CDR1 of the heavy chain of antibodies Act-1 and

; OTHER INFORMATION: LDP-02

US-09-748-960-12

Query Match 69.7%; Score 23; DB 10; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWM 4  
|||  
Db 2 YWM 4

## RESULT 12

US-09-832-312-49

; Sequence 49, Application US/09832312

; Patent No. US20010049829A1  
; GENERAL INFORMATION:  
; APPLICANT: Busfield et al.  
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF  
; FILE REFERENCE: 7853-234  
; CURRENT APPLICATION NUMBER: US/09/832,312  
; CURRENT FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: 09/610,118  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 09/503,387  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 09/454,824  
; PRIOR FILING DATE: 1999-12-06  
; PRIOR APPLICATION NUMBER: 09/345,468  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 49  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-832-312-49

Query Match 69.7%; Score 23; DB 10; Length 5;  
Best Local Similarity 75.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
|||  
Db 2 YWIS 5

## RESULT 13

US-09-770-002-1

; Sequence 1, Application US/09770002

; Patent No. US20020110558A1

; GENERAL INFORMATION:

; APPLICANT: Peter Lloyd Amlot

; APPLICANT: Max H. Schreier

; APPLICANT: Karin Schreier

; TITLE OF INVENTION: Use of CD25 binding molecules in the

; FILE REFERENCE: 4-30583A/30967C1

; CURRENT APPLICATION NUMBER: US/09/770,002

; CURRENT FILING DATE: 2001-09-17

; PRIOR APPLICATION NUMBER: PCT/EP99/05316

; PRIOR FILING DATE: 1999-07-26

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 5

; TYPE: PRT

; ORGANISM: mus musculus

; FEATURE:

US-09-770-002-1

Query Match 69.7%; Score 23; DB 10; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWM 4  
|||  
Db 2 YWM 4

## RESULT 14

US-09-291-417-40

; Sequence 40, Application US/09291417A

; Publication No. US20030050230A1

; GENERAL INFORMATION:

; APPLICANT: FLOWMAN, GREGORY

; APPLICANT: MARTINEZ, RICARDO

; APPLICANT: WHYTE, DAVID

; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; FILE REFERENCE: 240/300  
; CURRENT APPLICATION NUMBER: US/09/291,417A  
; CURRENT FILING DATE: 1999-04-13  
; EARLIER APPLICATION NUMBER: US 60/081,784  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 147  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 40  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: CCK4b  
US-09-291-417-40

Query Match 60.6%; Score 20; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WMS 5  
|||  
Db 1 WMS 3

RESULT 15  
US-10-029-301-4  
; Sequence 4, Application US/10029301  
; Publication No. US20020188107A1  
; GENERAL INFORMATION:  
; APPLICANT: Mynott, Tracy Lehanne  
; APPLICANT: Engwerda, Christian  
; APPLICANT: Peak, Keith  
; TITLE OF INVENTION: Component of Stem Bromelain (as amended)  
; FILE REFERENCE: 0623.0760002  
; CURRENT APPLICATION NUMBER: US/10/029,301  
; CURRENT FILING DATE: 2001-12-28  
; PRIOR APPLICATION NUMBER: US 09/382,685  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: PCT/GB98/00592  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: GB 9704252.7  
; PRIOR FILING DATE: 1997-02-28  
; PRIOR APPLICATION NUMBER: GB 9703850.9  
; PRIOR FILING DATE: 1997-02-25  
; PRIOR APPLICATION NUMBER: GB 9703827.7  
; PRIOR FILING DATE: 1997-02-25  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Ananas comosus  
US-10-029-301-4

Query Match 57.6%; Score 19; DB 9; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.7e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YWM 4  
|||  
Db 1 YWI 3

Search completed: April 22, 2003, 13:47:05  
Job time : 16 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:39:17 ; Search time 28 Seconds  
(without alignments)  
5.254 Million cell updates/sec

Title: US-09-674-716B-9

Perfect score: 33

Sequence: 1 GYVMS 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 19691

Minimum DB seq length: 0

Maximum DB seq length: 5

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	87.9	5	2	US-08-480-434-22
2	29	87.9	5	2	US-08-480-434-31
3	29	87.9	5	2	US-08-053-451B-22
4	29	87.9	5	2	US-08-053-451B-31
5	27	81.8	5	1	US-08-244-626-12
6	27	81.8	5	2	US-08-480-434-32
7	27	81.8	5	2	US-08-318-157B-23
8	27	81.8	5	2	US-08-053-451B-32
9	25	75.8	5	1	US-08-353-400-27
10	24	72.7	5	6	5185431-15
11	23	69.7	5	1	US-08-318-970B-1
12	23	69.7	5	1	US-08-318-970B-3
13	23	69.7	5	4	US-08-406-532-5
14	23	69.7	5	4	US-08-479-089A-7
15	20	60.6	4	1	US-07-869-933-21
16	20	60.6	4	4	US-08-103-663-21
17	20	60.6	5	2	US-08-476-176B-50
18	20	60.6	5	3	US-08-127-721A-50
19	20	60.6	5	3	US-08-485-246A-50
20	20	60.6	5	6	5217869-8
21	19	57.6	5	3	US-08-469-141A-64
22	19	57.6	5	4	US-07-987-264-1
23	19	57.6	5	4	US-09-382-689A-4
24	19	57.6	5	5	PCT-US95-13794-64
25	19	57.6	5	6	5217869-44
26	18	54.5	3	1	US-08-440-504A-1
27	18	54.5	3	4	US-08-433-613-1

28	18	54.5	4	1	US-08-321-625-14	Sequence 14, Appl
29	18	54.5	4	1	US-08-222-851-19	Sequence 19, Appl
30	18	54.5	4	1	US-08-469-486-39	Sequence 39, Appl
31	18	54.5	4	1	US-08-440-504A-2	Sequence 2, Appl
32	18	54.5	4	2	US-08-469-658-39	Sequence 39, Appl
33	18	54.5	4	2	US-08-671-487A-8	Sequence 8, Appl
34	18	54.5	4	4	US-08-651-179B-2	Sequence 2, Appl
35	18	54.5	4	4	US-08-433-613-2	Sequence 2, Appl
36	18	54.5	4	4	US-09-181-083-14	Sequence 14, Appl
37	18	54.5	4	4	US-09-155-613A-31	Sequence 31, Appl
38	18	54.5	4	4	US-09-441-063-1	Sequence 1, Appl
39	18	54.5	5	1	US-08-068-947-1	Sequence 1, Appl
40	18	54.5	5	1	US-08-325-509-50	Sequence 50, Appl
41	18	54.5	5	1	US-08-332-071B-1	Sequence 1, Appl
42	18	54.5	5	1	US-07-946-421-4	Sequence 4, Appl
43	18	54.5	5	1	US-08-321-625-3	Sequence 3, Appl
44	18	54.5	5	1	US-08-321-625-15	Sequence 15, Appl
45	18	54.5	5	1	US-07-789-184-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1  
US-08-480-434-22  
; Sequence 22, Application US/08480434  
; Patent No. 5811248  
; GENERAL INFORMATION:  
; APPLICANT: Charles C. Dittlow, et al.  
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC ANTIGENS,  
; TITLE OF INVENTION: ANTIBODIES THERETO, AND USES THEREOF  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,434  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Albert P. Halluin  
; REGISTRATION NUMBER: 25,227  
; REFERENCE/DOCKET NUMBER: 7606-053  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 854-3660  
; TELEFAX: (415) 854-3694  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: N  
; ANTI-SENSE: N  
; US-08-480-434-22

Query Match 87.9%; Score 29; DB 2; Length 5;  
Best Local Similarity 80.0%; Pred No. 2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYVMS 5  
|:|

Db 1 GFWS 5

## RESULT 2

US-08-480-434-31  
; Sequence 31, Application US/08480434  
; Patent No. 5811248

## GENERAL INFORMATION:

; APPLICANT: Charles C. Dittow, et al.  
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC ANTIGENS,  
; TITLE OF INVENTION: ANTIBODIES THERETO, AND USES THEREOF  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,434  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

; NAME: Albert P. Halluin  
; REGISTRATION NUMBER: 25,227  
; REFERENCE/DOCKET NUMBER: 7606-053  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 854-3660  
; TELEFAX: (415) 854-3694

## INFORMATION FOR SEQ ID NO: 31:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: N

; ANTI-SENSE: N

US-08-480-434-31

## Query Match

Best Local Similarity 87.9%; Score 29; DB 2; Length 5;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFWS 5

Db 1 GFWS 5

## RESULT 3

US-08-053-451B-22

; Sequence 22, Application US/08053451B  
; Patent No. 5955584

## GENERAL INFORMATION:

; APPLICANT: Chen, Francis W.  
; APPLICANT: Dittow, Charles C.  
; APPLICANT: Calenoff, Emanuel

; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC

; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF

; NUMBER OF SEQUENCES: 176

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA

; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/053,451B  
; FILING DATE: 26-APR-1993  
; CLASSIFICATION: 424

## ATTORNEY/AGENT INFORMATION:

; NAME: Halluin, Albert P.  
; REGISTRATION NUMBER: 25,227  
; REFERENCE/DOCKET NUMBER: 7606-033-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-854-3660  
; TELEFAX: 415-854-3694  
; TELEX: 66141 PENNIE

## INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA

; HYPOTHETICAL: N

; ANTI-SENSE: N

US-08-053-451B-22

## Query Match

Best Local Similarity 87.9%; Score 29; DB 2; Length 5;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFWS 5

Db 1 GFWS 5

## RESULT 4

US-08-053-451B-31

; Sequence 31, Application US/08053451B  
; Patent No. 5955584

## GENERAL INFORMATION:

; APPLICANT: Chen, Francis W.

; APPLICANT: Dittow, Charles C.

; APPLICANT: Calenoff, Emanuel

; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC

; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF

; NUMBER OF SEQUENCES: 176

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/053,451B

; FILING DATE: 26-APR-1993

; CLASSIFICATION: 424

## ATTORNEY/AGENT INFORMATION:

; NAME: Halluin, Albert P.

; REGISTRATION NUMBER: 25,227

; REFERENCE/DOCKET NUMBER: 7606-033-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-854-3660

; TELEFAX: 415-854-3694

; TELEX: 66141 PENNIE



INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
HYPOTHETICAL: N  
ANTI-SENSE: N  
US-08-053-451B-31

Query Match 87.9%; Score 29; DB 2; Length 5;  
Best Local Similarity 80.0%; Pred. No. 2e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWS 5  
|:|  
DB 1 GYWS 5

RESULT 5  
US-08-244-626-12  
Sequence 12, Application US/08244626  
Patent No. 5502167  
GENERAL INFORMATION:  
APPLICANT: Waldmann, Herman  
APPLICANT: Walsh, Louise  
APPLICANT: Crowe, James Scott  
APPLICANT: Lewis, Alan Peter  
TITLE OF INVENTION: CDR GRAFTED HUMANISED CHIMERIC T-CELL  
TITLE OF INVENTION: ANTIBODIES  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,626  
FILING DATE: July 15, 1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/02251  
FILING DATE: December 4, 1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Ernst, Barbara G.  
REGISTRATION NUMBER: 30,377  
REFERENCE/DOCKET NUMBER: 1808-153A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 783-6040  
TELEFAX: (202) 783-6031

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-244-626-12

Query Match 81.8%; Score 27; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
|:|  
DB 2 YWMS 5

Db 2 YWMS 5

RESULT 6  
US-08-480-434-32  
Sequence 32, Application US/08480434  
Patent No. 5811248  
GENERAL INFORMATION:  
APPLICANT: Charles C. Ditlow, et al.  
TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC ANTIGENS,  
ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,434  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Albert P. Halluin  
REGISTRATION NUMBER: 25,227  
REFERENCE/DOCKET NUMBER: 7606-053  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 854-3660  
TELEFAX: (415) 854-3694  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: N  
ANTI-SENSE: N  
US-08-480-434-32

Query Match 81.8%; Score 27; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5  
|:|  
DB 2 YWMS 5

RESULT 7  
US-08-318-157B-23  
Sequence 23, Application US/08318157B  
Patent No. 5874540  
GENERAL INFORMATION:  
APPLICANT: HANSEN, Hans J.  
APPLICANT: ARMOUR, Kathryn L.  
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
MOUSE MONOCLONAL ANTIBODIES  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

```
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/318,157B
/ FILING DATE: 05-OCT-1994
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Saxe, Bernhard D.
/ REGISTRATION NUMBER: 28,665
/ REFERENCE/DOCKET NUMBER: 18733/464
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202)672-5300
/ TELEFAX: (202)672-5399
/ TELEX: 904136
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-318-157B-23

Query Match      81.8%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 YWMS 5
Db  2 YWMS 5

RESULT 8
US-08-053-451B-32
; Sequence 32, Application US/08053451B
; Patent No. 595584
; GENERAL INFORMATION:
; APPLICANT: Chen, Francis W.
; APPLICANT: Ditlow, Charles C.
; APPLICANT: Calenoff, Emanuel
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC
; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/053,451B
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 7606-033-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
```

```
/
/ TYPE: amino acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA
/ HYPOTHETICAL: N
/ ANTI-SENSE: N
US-08-053-451B-32

Query Match      81.8%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 YWMS 5
Db  2 YWMS 5

RESULT 9
US-08-353-400-27
; Sequence 27, Application US/08353400
; Patent No. 5665357
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 37
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,400
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9324819.3
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411089.7
; FILING DATE: 03-JUN-1994
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-353-400-27

Query Match      75.8%; Score 25; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 2e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 GYWM 4
Db  1 GYWI 4

RESULT 10
5185431-15
; Patent No. 5185431
; APPLICANT: YOSHIMATSU, KENTARO;SHIKATA, YASUSHI;TANAKA, ISAO;
; HASEGAWA, YOSHIKAZU;SETO, TOSHIO;OSAWA, TOSHIO
; TITLE OF INVENTION: RECOMBINANT NATURAL KILLER CELL ACTIVATOR
; NUMBER OF SEQUENCES: 31
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/392,841
; FILING DATE: 11-AUG-1989
; SEQ ID NO:15:
; LENGTH: 5
5185431-15

Query Match      72.7%; Score 24; DB 6; Length 5;
```

```

Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

Qy 1 GYM 3
Db 2 GYM 4

RESULT 11
US-08-318-970B-1
; Sequence 1, Application US/08318970B
; Patent No. 5589573
; GENERAL INFORMATION:
; APPLICANT: Hideaki HAGIWARA, et al.
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF ANTI-IDIOTYPIC
; TITLE OF INVENTION: ANTIBODIES AGAINST ANTI-CANCER HUMAN MONOCLONAL ANTIBODY
; TITLE OF INVENTION: AND DNA BASE SEQUENCES ENCODING THOSE SEQUENCES
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Office of Sherman and Shalloway
; STREET: 413 N. Washington Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: Dell System 210; Intel 80 285 Microprocessor
; OPERATING SYSTEM: MS DOS 3.3
; SOFTWARE: Word Perfect, Version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,970B
; FILING DATE: October 6, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard A. Steinberg
; REGISTRATION/DOCKET NUMBER: S-2371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 549-2282
; TELEFAX: (703) 836-0106
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: H-CDRI-1
; OTHER INFORMATION: hypervariable region
US-08-318-970B-1

Query Match 69.7%; Score 23; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

Qy 2 YWM 4
Db 2 YWM 4

RESULT 12
US-08-318-970B-3
; Sequence 3, Application US/08318970B
; Patent No. 5589573
; GENERAL INFORMATION:
; APPLICANT: Hideaki HAGIWARA, et al.
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF ANTI-IDIOTYPIC
; TITLE OF INVENTION: ANTIBODIES AGAINST ANTI-CANCER HUMAN MONOCLONAL ANTIBODY
; TITLE OF INVENTION: AND DNA BASE SEQUENCES ENCODING THOSE SEQUENCES
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Office of Sherman and Shalloway

```

```

; STREET: 413 N. Washington Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: Dell System 210; Intel 80 285 Microprocessor
; OPERATING SYSTEM: MS DOS 3.3
; SOFTWARE: Word Perfect, Version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,970B
; FILING DATE: October 6, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard A. Steinberg
; REGISTRATION/DOCKET NUMBER: 26,588
; REFERENCE/DOCKET NUMBER: S-2371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 549-2282
; TELEFAX: (703) 836-0106
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: H-CDRI-3
; OTHER INFORMATION: hypervariable region
US-08-318-970B-3

Query Match 69.7%; Score 23; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

Qy 2 YWM 4
Db 2 YWM 4

RESULT 13
US-09-406-532-5
; Sequence 5, Application US/09406532A
; Patent No. 6365154
; GENERAL INFORMATION:
; APPLICANT: Connie L. Erickson-Miller
; APPLICANT: Stephen D. Holmes
; APPLICANT: James D. Winkler
; TITLE OF INVENTION: TIE2 Agonist Antibodies
; FILE REFERENCE: P50843
; CURRENT APPLICATION NUMBER: US/09/406,532A
; CURRENT FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 60/102,098
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)...(5)
; OTHER INFORMATION: 1588 heavy chain CDR 1
US-09-406-532-5

Query Match 69.7%; Score 23; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

Qy 2 YWM 4
Db 2 YWM 4

```

```
Db      2 YWM 4

RESULT 14
US-08-479-089A-7
; Sequence 7, Application US/08479089A
; Patent No. 6383487
; GENERAL INFORMATION:
; APPLICANT: Amlot, Peter L.
; APPLICANT: Akbar, Arne N.
; APPLICANT: Heinrich, Gunther
; APPLICANT: Cammisuli, Salvatore
; TITLE OF INVENTION: CD25 Binding Molecules
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 638487artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: NJ
; COUNTRY: USA
; ZIP: 07901-1027
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,089A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/669,545
; FILING DATE:
; APPLICATION NUMBER: GB 9019323
; FILING DATE: 05-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Furman, Diane E.
; REGISTRATION NUMBER: 31,104
; REFERENCE/DOCKET NUMBER: 100-7617
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 522-6924
; TELEFAX: (908) 522-6955
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
;
US-08-479-089A-7

Query Match      69.7%; Score 23; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 YWM 4
Db      2 YWM 4

RESULT 15
US-07-869-933-21
; Sequence 21, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-07-869-933-21

Query Match      60.6%; Score 20; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 2e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 YWM 4
Db      2 YWL 4

Search completed: April 22, 2003, 13:43:03
Job time : 29 secs
```

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:46:48 ; Search time 14 seconds  
(without alignments)  
130.468 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98

Sequence: 1 EIRLSDNYATHVAESVKG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 3433

Minimum DB seq length: 0  
Maximum DB seq length: 19

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26.5	27.0	15	2 A39109	hypothetical prote
2	25	25.5	10	2 S77990	cytochrome-c oxida
3	24	24.5	19	2 S39387	protein kinase GSK
4	23	23.5	15	2 PH1631	Ig H chain V-D-J r
5	23	23.5	16	2 B60278	24k antigen - Myco
6	22	22.4	10	2 D46278	formaldehyde dehyd
7	22	22.4	14	2 S59435	formate dehydrogen
8	22	22.4	15	2 PA0046	protein QA100044 -
9	22	22.4	15	2 I67525	CD33 antigen homol
10	22	22.4	18	2 F49215	urease (EC 3.5.1.5
11	22	22.4	18	2 PQ0072	T-cell receptor be
12	22	22.4	19	2 A41077	protein-disulfide
13	21	21.4	9	2 S30494	cat gene leader pe
14	21	21.4	9	2 B24362	chloramphenicol O-
15	21	21.4	12	2 PH1587	Ig H chain V-D-J r
16	21	21.4	14	2 PH1471	T-cell receptor be
17	21	21.4	16	2 A29520	amino-acid racemas
18	21	21.4	17	2 A61334	trypsin (EC 3.4.21
19	21	21.4	18	2 PH0768	T-cell receptor be
20	21	21.4	18	4 I56393	lacZ/IS1 mutant fu
21	20	20.4	9	2 PT0238	Ig heavy chain CRD
22	20	20.4	10	2 S65388	cytochrome-c oxida
23	20	20.4	11	2 PT0301	Ig heavy chain CRD
24	20	20.4	12	2 S26546	T-cell receptor be
25	20	20.4	12	4 PC2122	aminotransferase c
26	20	20.4	13	2 PH1595	Ig H chain V-D-J r
27	20	20.4	14	2 A47421	leukotriene B-4 12
28	20	20.4	15	2 A30330	neuropeptide pep -
29	20	20.4	15	2 I78838	flt3 ligand isofor

30 20 20.4 16 2 S17217 ribulose-bisphosph  
31 20 20.4 16 2 S51610 hypothetical prote  
32 20 20.4 16 2 PH1622 Ig H chain V-D-J r  
33 20 20.4 16 2 PH1589 Ig H chain V-D-J r  
34 20 20.4 17 2 C85956 hypothetical prote  
35 20 20.4 18 2 PH1621 Ig H chain V-D-J r  
36 20 20.4 18 2 A59396 Tha p 1 - Thaumeto  
37 19.5 19.9 16 2 D58501 26k kidney and gal  
38 19 19.4 10 1 GMROL2 leucosulfakinin-II  
39 19 19.4 10 2 B61033 ranachykinin B -  
40 19 19.4 10 2 B60656 leucosulfakinin II  
41 19 19.4 12 2 G40003 hypothetical prote  
42 19 19.4 12 2 D28551 hypothetical prote  
43 19 19.4 12 2 A34858 proteinase E - bla  
44 19 19.4 12 2 C30503 proteinase-2b chain  
45 19 19.4 13 2 PN0123 serine proteinase

#### ALIGNMENTS

##### RESULT 1

A39109  
hypothetical protein 1 - hepatitis C virus  
C:Species: hepatitis C virus  
C:Date: 18-Oct-1991 #sequence\_revision 18-Oct-1991 #text\_change 18-Jun-1993  
C:Accession: A39109  
R:Han, J.H.; Shyamala, V.; Richman, K.H.; Brauer, M.J.; Irvine, B.; Urdea, M.S.; Tekamp-  
Proc. Natl. Acad. Sci. U.S.A. 88, 1711-1715, 1991  
A:Title: Characterization of the terminal regions of hepatitis C viral RNA: identificati.  
A:Reference number: A39109; MUID:91156678; PMID:1705704  
A:Accession: A39109  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-16 <HAN>  
A:Cross-references: GB:M58406

Query Match 27.0%; Score 26.5; DB 2; Length 16;  
Best Local Similarity 70.0%; Pred. No. 5.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 8 NVATHVAESV 17  
||| |||||  
DB 8 NYCLH-AESV 16

##### RESULT 2

S77990  
cytochrome-c oxidase (EC 1.9.3.1) chain VIIc - bigeye tuna (fragment)  
C:Species: Thunnus obsesus (bigeye tuna)  
C:Date: 17-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 30-Jan-1998  
C:Accession: S77990  
R:Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.  
submitted to the Protein Sequence Database, June 1997  
A:Reference number: S77980  
A:Accession: S77990  
A:Molecule type: protein  
A:Residues: 1-10 <ARN>  
A:Experimental source: heart; liver  
C:Genetics:  
A:Genome: nuclear  
C:Function:  
A:Pathway: oxidative phosphorylation; respiratory chain  
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 25.5%; Score 25; DB 2; Length 10;  
Best Local Similarity 80.0%; Pred. No. 5.6e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 THYAE 15  
:|||||  
DB 1 SHYAE 5

## RESULT 3

S39387  
 protein kinase GSK-3-beta (EC 2.7.1.-) - rabbit (fragment)  
 C;Species: Oryctolagus cuniculus (domestic rabbit)  
 C;Date: 19-May-1994 #sequence\_revision 26-Jul-1996 #text\_change 07-May-1999  
 C;Accession: S39387  
 R;Sutherland, C.; Leighton, I.A.; Cohen, P.  
 Biochem. J. 296, 15-19, 1993  
 A;Title: Inactivation of glycogen synthase kinase-3-beta by phosphorylation: new kinase  
 A;Reference number: S39387; MUID:94071817; PMID:8250835  
 A;Accession: S39387  
 A;Molecule type: protein  
 A;Residues: 1-19 <SVT>  
 A;Status: preliminary  
 C;Superfamily: kinase-related transforming protein; protein kinase homology  
 C;Keywords: phosphotransferase; protein kinase

Query Match 24.5%; Score 24; DB 2; Length 19;  
 Best Local Similarity 62.5%; Pred. No. 1.6e+03;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 THYAESVK 18  
 | : || |  
 Db 2 TSFAESXK 9

## RESULT 4

PH1631  
 Ig H chain V-D-J region (clone B-less 202) - mouse (fragment)  
 C;Species: Mus musculus (house mouse)  
 C;Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999  
 C;Accession: PH1631  
 R;Levinson, D.A.; Campos-Torres, J.; Leder, P.  
 J. Exp. Med. 178, 317-329, 1993  
 A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice  
 A;Reference number: PH1580; MUID:93301609; PMID:8315387  
 A;Accession: PH1631  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LEV>  
 A;Experimental source: bone marrow pre-B lymphocyte  
 C;Keywords: immunoglobulin

Query Match 23.5%; Score 23; DB 2; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 SDNYATHY 13  
 | : || |  
 Db 8 TDGYAMDY 15

## RESULT 5

B60278  
 24K antigen - Mycobacterium bovis (fragment)  
 C;Species: Mycobacterium bovis  
 C;Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 18-Jun-1993  
 C;Accession: B60278  
 R;Fitis, T.; Costopoulos, C.; Radford, A.J.; Bacic, A.; Wood, P.R.  
 Infect. Immun. 59, 800-807, 1991  
 A;Title: Purification and characterization of major antigens from a Mycobacterium bovis  
 A;Reference number: A60278; MUID:91147217; PMID:1900061  
 A;Accession: B60278  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-16 <FIP>

Query Match 23.5%; Score 23; DB 2; Length 16;  
 Best Local Similarity 57.1%; Pred. No. 2e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 YAESVK 19  
 | : || |  
 Db 5 YKELKG 11

## RESULT 6

D46285  
 formaldehyde dehydrogenase (glutathione) (EC 1.2.1.1) class III low activity form - Atlantic  
 C;Species: Gadus morhua (Atlantic cod)  
 C;Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1996  
 C;Accession: D46285  
 R;Danielsson, O.; Jornvall, H.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 9247-9251, 1992  
 A;Title: "Enzymogenesis": classical liver alcohol dehydrogenase origin from the glutathione  
 A;Reference number: A46285; MUID:93028441; PMID:1409630  
 A;Accession: D46285  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-10 <DAN>  
 A;Note: sequence extracted from NCBI backbone (NCBI:116272)  
 C;Keywords: NAD; oxidoreductase

Query Match 22.4%; Score 22; DB 2; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DNYATH 12  
 | : || |  
 Db 1 DEFVTH 6

## RESULT 7

S59495  
 formaldehyde dehydrogenase delta chain - Alcaligenes eutrophus (fragment)  
 C;Species: Alcaligenes eutrophus  
 C;Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999  
 C;Accession: S59495  
 R;Friedebold, J.; Mayer, F.; Bill, E.; Trautwein, A.X.; Bowien, B.  
 Biol. Chem. Hoppe-Seyler 376, 561-568, 1995  
 A;Title: Structural and immunological studies on the soluble formaldehyde dehydrogenase from  
 A;Reference number: S59492; MUID:96145736; PMID:8561915  
 A;Accession: S59495  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-14 <FRI>

Query Match 22.4%; Score 22; DB 2; Length 14;  
 Best Local Similarity 44.4%; Pred. No. 2.5e+03;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 LKSDNYATH 12  
 | : || |  
 Db 1 MKIDNIITY 9

## RESULT 8

PA0046  
 protein QA100044 - Arabidopsis thaliana (fragment)  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 06-Jun-1997  
 C;Accession: PA0046; PA0042  
 R;Kamo, M.; Kawakami, T.; Miyatake, N.; Taugita, A.  
 submitted to JIPID July 1994  
 A;Description: Separation and characterization of Arabidopsis proteins by two-dimensions  
 A;Reference number: PA0001  
 A;Accession: PA0046  
 A;Molecule type: protein  
 A;Residues: 1-15 <KAM>  
 A;Experimental source: stem

Query Match 22.4%; Score 22; DB 2; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 2.7e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 AESVKG 19  
 | : || |

Db 1 AESKKG 6

## RESULT 9

167525  
CD33 antigen homolog - mouse (fragment)  
C:Species: Mus sp. (mouse)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Jun-1998  
C:Accession: I67525  
R:Chies, J.A.; Lembezat, M.P.; Freitas, A.A.  
Eur. J. Immunol. 24, 1657-1664, 1994  
A:Title: Entry of B lymphocytes into the persistent cell pool in non-immunized mice is regulated by CD33  
A:Reference number: 153392; MUID:94298870; PMID:8026526  
A:Accession: I67525  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-15 <RES>  
A:Cross-references: GB:S71349; NID:9550037  
C:Genetics:  
A:Gene: IG VH7183

Query Match 22.4%; Score 22; DB 2; Length 15;  
Best Local Similarity 33.3%; Pred. No. 2.7e+03;  
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 KSDNYATHY 13  
: ||| :  
Db 4 RRDHYGSSY 12

## RESULT 10

F49215  
urease (EC 3.5.1.5) small chain UreA - Helicobacter mustelae (ATCC 43772) (fragment)  
C:Species: Helicobacter mustelae  
C:Date: 19-Dec-1993 #sequence\_revision 25-Apr-1997 #text\_change 30-May-1997  
C:Accession: F49215  
R:Turbett, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.  
Infect. Immun. 60, 5259-5266, 1992  
A:Title: Purification and characterization of the urease enzymes of Helicobacter species  
A:Reference number: A49215; MUID:93084378; PMID:1452359  
A:Accession: F49215  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-18 <TUR>  
A:Note: sequence extracted from NCBI backbone (NCBIP:119487)  
C:Superfamily: urease 26K chain; urease 11K chain homology; urease 12K chain homology  
C:Keywords: hydrolase

Query Match 22.4%; Score 22; DB 2; Length 18;  
Best Local Similarity 50.0%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 DNYATHYA 14  
: ||| :  
Db 7 DKMMLHYA 14

## RESULT 11

PQ0072  
T-cell receptor beta chain (BTB98) - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Feb-1995  
C:Accession: PQ0072  
R:Tanaka, A.; Ishiguro, N.; Shinagawa, M.  
submitted to JIPID, May 1990  
A:Description: Sequence analysis of bovine T-cell receptor beta chain genes.  
A:Reference number: JQ0472  
A:Accession: PQ0072  
A:Molecule type: mRNA  
A:Residues: 1-18 <TAN>  
C:Genetics:  
A:Gene: BTB98

C:Keywords: receptor

Query Match 22.4%; Score 22; DB 2; Length 18;  
Best Local Similarity 38.5%; Pred. No. 3.2e+03;  
Matches 5; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

Qy 6 SDNYATHYAESVK 18  
: ||| :  
Db 4 SDDY--HFGPGTK 14

## RESULT 12

A41077  
protein-disulfide reductase (glutathione) (EC 1.8.4.2) Q-5 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 12-Jun-1992 #sequence\_revision 12-Jun-1992 #text\_change 23-Jun-1993  
C:Accession: A41077  
R:Srivastava, S.P.; Chen, N.; Liu, Y.; Holtzman, J.L.  
J. Biol. Chem. 266, 20337-20344, 1991  
A:Title: Purification and characterization of a new isozyme of thiol:protein-disulfide oxidoreductase from rat liver  
A:Reference number: A41077; MUID:92041865; PMID:1657921  
A:Accession: A41077  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-19 <SRI>  
C:Keywords: oxidoreductase

Query Match 22.4%; Score 22; DB 2; Length 19;  
Best Local Similarity 57.1%; Pred. No. 3.4e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LKSDNYA 10  
: ||| :  
Db 12 LKSNFA 18

## RESULT 13

S30494  
cat gene leader peptide - Streptococcus agalactiae plasmid pIP501  
C:Species: Streptococcus agalactiae  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 24-Sep-1999  
C:Accession: S30494  
R:Trieu-Cuot, P.; de Cespedes, G.; Horaud, T.  
plasmid 28, 272-276, 1992  
A:Title: Nucleotide sequence of the chloramphenicol resistance determinant of the Streptococcus agalactiae plasmid pIP501  
A:Reference number: JQ1950; MUID:93096867; PMID:1461942  
A:Accession: S30494  
A:Molecule type: DNA  
A:Residues: 1-9 <TRI>  
A:Cross-references: EMBL:X65462; NID:949071; PIDN:CAA46454.1; PID:G581554  
C:Genetics:  
A:Genome: plasmid pIP501  
C:Superfamily: unassigned leader peptides

Query Match 21.4%; Score 21; DB 2; Length 9;  
Best Local Similarity 42.9%; Pred. No. 2.8e+05;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KSDNYAT 11  
: ||| :  
Db 3 KSEDYSS 9

## RESULT 14

B24362  
chloramphenicol O-acetyltransferase leader peptide - Staphylococcus aureus plasmid PUB11  
C:Species: Staphylococcus aureus  
C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 11-May-2000  
C:Accession: B24362  
R:Bruckner, R.; Matzura, H.  
EMBO J. 4, 2295-2300, 1985  
A:Title: Regulation of the inducible chloramphenicol acetyltransferase gene of the Staphylococcus aureus plasmid PUB11

A;Reference number: A24362; MUID:86081739; PMID:3865770  
 A;Accession: B24362  
 A;Molecule type: DNA  
 A;Residues: 1-9 <BRU>  
 A;Cross-references: GB:X02872; NID:946536; PIDN:CAA26630.1; PID:9581555  
 C;Comment: Ribosome stalling in the translation of this leader peptide, caused by the ef  
 relation of the chloramphenicol O-acetyltransferase from a ribosome binding site located  
 C;Genetics:  
 A;Genome: plasmid  
 C;Superfamily: unassigned leader peptides

Query Match 21.4%; Score 21; DB 2; Length 9;  
 Best Local Similarity 42.9%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 5 KSDNYAT 11  
 ||:|:  
 Db 3 KSEDYSS 9

RESULT 15  
 PH1587  
 Ig H chain V-D-J region (wild-type clone 11) - mouse (fragment)  
 C;Species: Mus musculus (house mouse)  
 C;Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999  
 C;Accession: PH1587  
 R;Levinson, D.A.; Campos-Torres, J.; Leder, P.  
 J. Exp. Med. 178, 317-329, 1993  
 A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice  
 A;Reference number: PH1580; MUID:93301609; PMID:8315387  
 A;Accession: PH1587  
 A;Molecule type: DNA  
 A;Residues: 1-12 <LEV>  
 A;Experimental source: bone marrow pre-B lymphocyte  
 C;Keywords: immunoglobulin

Query Match 21.4%; Score 21; DB 2; Length 12;  
 Best Local Similarity 44.4%; Pred. No. 3e+03;  
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 KSDNYATHY 13  
 :|||  
 Db 4 RGDLYANDY 12

Search completed: April 22, 2003, 13:49:54  
 Job time : 15 secs



GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:43:07 ; Search time 25 Seconds  
(without alignments)  
31.522 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98

Sequence: 1 EIRKSDNYATHVAESVKG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1015

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	27.6	14	1 NEJ2_FASHE	P80526 fasciola he
2	25	25.5	10	1 COXO_THUOB	P80982 thunnus obe
3	22	22.4	12	1 LOSK_LOCM1	P47733 locusta mig
4	21	21.4	9	1 LPCA_STRAU	P36884 staphylococ
5	21	21.4	16	1 ALRX_PSEPU	P17916 pseudomonas
6	21	21.4	18	1 RIP_SIRGR	P83323 siraitia gr
7	21	21.4	18	1 UC21_MAYZE	P80627 zea mays (m
8	20	20.4	10	1 COXO_RAT	P80432 rattus norv
9	20	20.4	17	1 AU32_LITRA	P82395 litoria ran
10	20	20.4	19	1 ATPB_CANFA	P99504 canis fami
11	19	19.4	10	1 LSK2_LEUMA	P09039 leucophaea
12	19	19.4	10	1 TKNB_RANCA	P22689 rana catesb
13	19	19.4	14	1 UN37_CLOPA	P81358 clostridium
14	19	19.4	15	1 MILT_ONCKE	P81037 oncorhynch
15	19	19.4	15	1 NIS1_ANAVA	Q44507 anabaena va
16	19	19.4	16	1 RL6_VIBPR	Q56715 vibrio prot
17	19	19.4	17	1 EFG_THEAQ	Q01697 thermus aqu
18	19	19.4	17	1 H170_RAT	P21794 rattus norv
19	18	18.4	12	1 TKN_FASSE	P08611 kassina sen
20	18	18.4	15	1 OBPA_MAMBR	P14285 mamestra br
21	18	18.4	15	1 YAA3_RHOPA	Q02006 rhodopsu
22	18	18.4	16	1 HTPG_ACICA	P81876 acinetobact
23	18	18.4	17	1 AU31_LITRA	P82394 litoria ran
24	18	18.4	17	1 B29K_PORGI	P81784 porphyromon
25	18	18.4	17	1 NU4M_TRIRU	Q36834 trichophyto
26	18	18.4	19	1 FIBB_LAMGL	P14473 lama glama
27	18	18.4	19	1 OXLA_OPHPA	P81383 ophiophagus
28	18	18.4	19	1 PHLC_STAIN	P80924 staphylococ
29	17	17.3	12	1 RR16_GINBI	P36207 ginkgo bilo
30	17	17.3	14	1 MAST_VESBA	P21654 vespa basal
31	17	17.3	15	1 MM01_RAT	P81563 rattus norv
32	17	17.3	17	1 SP51_BACLI	P27642 bacillus li
33	17	17.3	18	1 AROF_STRRM	P80575 streptomyc

34 17 17.3 19 1 LPRM\_STAAU P03063 staphylococ  
35 17 17.3 19 1 TX3\_PHONI P31010 phoneutria  
36 17 17.3 19 1 UP24\_UPEIN P82030 uperoleia i  
37 16 16.3 8 1 AL18\_CARMA P81821 carcinus ma  
38 16 16.3 8 1 RSI\_ERWCH P37985 erwinia chr  
39 16 16.3 10 1 GONI\_PETWA P04378 petromyzon  
40 16 16.3 10 1 TKNB\_CHICK P19851 gallus gall  
41 16 16.3 12 1 FARI\_CALVO P41869 calliphora  
42 16 16.3 12 1 PSP3\_PHYPA P80662 physcomitre  
43 16 16.3 13 1 AUL1\_LITRA P82386 litoria ran  
44 16 16.3 13 1 UP71\_LITWA P82050 litoria ewi  
45 16 16.3 15 1 CXAI\_CONGE P01519 conus geogr

## ALIGNMENTS

RESULT 1  
NEJ2\_FASHE  
ID NEJ2\_FASHE STANDARD; PRT; 14 AA.  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Newly excysted juvenile protein 2 (Fragment).  
OS Fasciola hepatica (liver fluke).  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
OC Echinostomida; Echinostomata; Fasciolidae; Fasciola.  
OX NCBI\_TaxID=6192;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95366993; PubMed=7639732;  
RA Tkalecivic J., Ashman K., Meusen E.;  
RT "Fasciola hepatica: rapid identification of newly excysted juvenile  
RT proteins";  
RL Biochem. Biophys. Res. Commun. 213:169-174 (1995).  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT THE NEWLY EXCYSTED JUVENILE  
STAGE.  
FT NON TER 14 14  
SQ SEQUENCE 14 AA; 1581 MW; 9E0F0090CC8C0DF1 CRC64;

Query Match 27.6%; Score 27; DB 1; Length 14;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DNYATHYA 14  
DB 3 DNGRTHWA 10

RESULT 2  
COXO\_THUOB  
ID COXO\_THUOB STANDARD; PRT; 10 AA.  
AC P80982;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Cytochrome c oxidase polypeptide VIIC (EC 1.9.3.1) (Fragment).  
OS Thunnus obesus (Bigeye tuna).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;  
OC Scombridae; Thunnus.  
OX NCBI\_TaxID=8241;  
RN [1]  
RP SEQUENCE.  
RX TISSUE=Heart, and Liver;  
RX MEDLINE=97454291; PubMed=9310366;  
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,  
RA Kadenbach B.;  
RT "The subunit structure of cytochrome-c oxidase from tuna heart and  
RT liver";  
RL Eur. J. Biochem. 248:99-103(1997).

CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE  
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN  
 CC MITOCHONDRIAL ELECTRON TRANSPORT.  
 CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.  
 CC Oxidoreductase; Inner membrane; Mitochondrion.  
 KW NON\_TER 10  
 FT 10  
 SQ SEQUENCE 10 AA; 1059 MW; 126DE76787B1DCB CRC64;

Query Match 25.5%; Score 25; DB 1; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 THYAE 15  
 DB 1 SHYAE 5

RESULT 3  
 LOSK\_LOCMI STANDARD; PRT; 12 AA.  
 AC P47733;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Sulfakinin (LOM-SK).  
 OS Locusta migratoria (Migratory locust).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 CC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
 CC Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]

RP SEQUENCE.  
 RC TISSUE=Brain;  
 RA Schoofs L., Holman G.L., Hayes T.K., Nachman R.J., de Loof A.;  
 RL (In) McCaffery A., Wilson I. (eds.);  
 RL Chromatography and Isolation of insect hormones and pheromones,  
 RL pp.231-241, Plenum Press, New York (1990).  
 CC -!- FUNCTION: MYOTROPIC PEPTIDE.  
 CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.  
 DR InterPro; IPR001651; Gastrin.  
 DR PROSITE; PS00259; GASTRIN.  
 KW Hormone; Amidation; Sulfation.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 7 7 SULFATION (POTENTIAL).  
 FT MOD\_RES 12 12 AMIDATION.  
 SQ SEQUENCE 12 AA; 1440 MW; 9B5B5DA9BD6B5AAA CRC64;

Query Match 22.4%; Score 22; DB 1; Length 12;  
 Best Local Similarity 57.1%; Pred. No. 6.7e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RLKSDNY 9  
 DB 1 QLASDDY 7

RESULT 4  
 LPGA\_STAAU STANDARD; PRT; 9 AA.  
 AC P36884;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Chloramphenicol resistance leader peptide.  
 OS Staphylococcus aureus, and  
 OS Streptococcus agalactiae.  
 CC Plasmid pSCS6, Plasmid pSCS7, Plasmid pUB112, and Plasmid pIP501.  
 CC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1280, 1311;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC SPECIES=S.aureus; STRAIN=436; PLASMID=pSCS7;  
 RX MEDLINE=92027652; PubMed=1929326;  
 RA Schwarz S., Cardoso M.;  
 RT "Nucleotide sequence and phylogeny of a chloramphenicol  
 RT acetyltransferase encoded by the plasmid pSCS7 from Staphylococcus  
 RT aureus.";  
 RL Antimicrob. Agents Chemother. 35:1551-1556(1991).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC SPECIES=S.aureus; PLASMID=pSCS6;  
 RX MEDLINE=92388047; PubMed=1517170;  
 RA Cardoso M., Schwarz S.;  
 RT "Nucleotide sequence and structural relationships of a  
 RT chloramphenicol acetyltransferase encoded by the plasmid pSCS6 from  
 RT Staphylococcus aureus.";  
 RL J. Appl. Bacteriol. 72:289-293(1992).  
 RN [3]

RP SEQUENCE FROM N.A.  
 RC SPECIES=S.aureus; PLASMID=pUB112;  
 RX MEDLINE=86081739; PubMed=3865770;  
 RA Brueckner R., Matzura H.;  
 RT "Regulation of the inducible chloramphenicol acetyltransferase gene  
 RT of the Staphylococcus aureus plasmid pUB112.";  
 RL EMBO J. 4:2295-2300(1985).  
 RN [4]

RP SEQUENCE FROM N.A.  
 RC SPECIES=S.agalactiae; PLASMID=PIP501;  
 RX MEDLINE=93096867; PubMed=1461942;  
 RA Trieu-Cuot P., de Cespedes G., Horaud T.;  
 RT "Nucleotide sequence of the chloramphenicol resistance determinant of  
 RT the streptococcal plasmid pIP501.";  
 RL Plasmid 28:272-276(1992).  
 RN [5]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M58515; AAA26612.1; -;  
 DR EMBL; M58516; AAA16528.1; -;  
 DR EMBL; X02872; CAA26630.1; -;  
 DR EMBL; X60827; CAA43217.1; -;  
 DR EMBL; X65462; CAA46454.1; -;  
 DR PIR; B24362; B24362  
 KW Leader peptide; Antibiotic resistance; Plasmid.  
 SQ SEQUENCE 9 AA; 1074 MW; 5D9CAB5AAB05B333 CRC64;

Query Match 21.4%; Score 21; DB 1; Length 9;  
 Best Local Similarity 42.9%; Pred. No. 1.1e+05;  
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 5 KSDNYAT 11  
 DB 3 KSEDYSS 9

RESULT 5  
 ALRX\_PSEPU STANDARD; PRT; 16 AA.  
 ID ALRX\_PSEPU  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Broad-specificity amino acid racemase (EC 5.1.1.-) (Fragment).  
 OS Pseudomonas putida.  
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas.  
 OX NCBI\_TaxID=303;  
 RN [1]

```

RP SEQUENCE.
RC STRAIN=AKU 0813;
RX MEDLINE=85072810; PubMed=6439237;
RA Roise D., Soda K., Yagi T., Walsh C.T.;
RT "Inactivation of the Pseudomonas striata broad specificity amino acid
RT racemase by D and L isomers of beta-substituted alanines: Kinetics,
RT stoichiometry, active site peptide, and mechanistic studies.";
RL Biochemistry 23:5195-5201(1984).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE ALANINE RACEMASE FAMILY.
DR PIR: A29520; A29520.
DR PROSITE; PS00395; ALANINE RACEMASE; 1.
KW Isomerase; Pyridoxal phosphate.
FT ACT_SITE 6 6
FT BINDING 6 6
FT NON_TER 16 16
FT PYRIDOXAL PHOSPHATE.
SQ SEQUENCE 16 AA; 1572 MW; 8C6C2DA154362E07 CRC64;

Query Match 21.4%; Score 21; DB 1; Length 16;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LKSDNY 9
DB 5 LKADAY 10

RESULT 6
ID RIP_SIRGR STANDARD; PRT; 18 AA.
AC P83323;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein momorgosvin (rRNA N-glycosidase) (EC
DE 3.2.2.22) (Fragment).
OS Siraitia grosvenorii (Monk's fruit) (Luo han guo).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosidis I; Cucurbitales; Cucurbitaceae; Siraitia.
OX NCBI_TaxID=190515;
RN [1]
RP SEQUENCE, FUNCTION, AND GLYCOSYLATION.
RC TISSUE=Seed;
RX MEDLINE=21073786; PubMed=11205869;
RA Tsang K.Y., Ng T.B.;
RT "Isolation and characterization of a new ribosome inactivating
RT protein, momorgosvin, from seeds of the monk's fruit Momordica
RT grosvenorii.";
RL Life Sci. 68:773-784(2001).
CC -1- FUNCTION: Inhibits protein synthesis in the rabbit reticulocyte
CC lysate system.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
DR InterPro; IPR001574; RIP.
DR PROSITE; PS00275; SHIGA_RICIN; PARTIAL.
KW Plant defense; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Glycoprotein.
FT NON_TER 18 18
FT SEQUENCE 18 AA; 2032 MW; 476E15CE683E2446 CRC64;

Query Match 21.4%; Score 21; DB 1; Length 18;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 NYATHY 13
DB 10 NGATY 15

```

```

RESULT 7
UC21_MAIZE
ID UC21_MAIZE STANDARD; PRT; 18 AA.
AC P80627;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 443)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.9, ITS MW IS: 26.5 kDa.
CC -1- SIMILARITY: TO L-ASCORBATE PEROXIDASES.
DR Maize-2DPAGE; P80627; COLEOPTILE.
DR MaizeDB; 123953; -.
FT NON_TER 1 1
FT NON_TER 18 18
FT SEQUENCE 18 AA; 1938 MW; F32F6FEF038BAB8A CRC64;

Query Match 21.4%; Score 21; DB 1; Length 18;
Best Local Similarity 62.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 NYATHYAE 15
DB 3 NYPTVSAE 10

RESULT 8
COXO_RAT
ID COXO_RAT STANDARD; PRT; 10 AA.
AC P80432;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide Viic, mitochondrial (EC 1.9.3.1)
DE (VIIIA) (Fragment).
GN COX7C OR COX7C1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar; TISSUE=Liver, and Heart;
RX MEDLINE=95324529; PubMed=7601105;
RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;
RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and
RT amino-terminal sequences suggest identity of the fetal heart and the
RT adult liver isoform.";
RL Eur. J. Biochem. 230:235-241(1995).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 10 10

```

```
SQ SEQUENCE 10 AA; 1117 MW; 126DE767687B1B0B CRC64;
Query Match 20.4%; Score 20; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 THYAE 15
Db 1 SHYEE 5

RESULT 9
AUS2_LITRA
ID AU32_LITRA STANDARD; PRT; 17 AA.
AC P82395;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aurein 3.2.
OS Litoria raniformis (Southern bell frog), and
OS Litoria aurea (Green and golden bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Pelodyadinae; Litoria.
OC NCBI_TaxID=116057, 8371;
RN [1]
SEQUENCE, AND FUNCTION.
RC TISSUE=Granular dorsal gland;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT Australian bell frogs Litoria aurea and Litoria raniformis the
RT solution structure of aurein 1.2."
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -1- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST L.LACTIS, L.INNOCUA,
CC M.LUTEUS, S.AUREUS, S.EPIDERMIDIS AND S.UBERTIS. PROBABLY ACTS BY
CC DISTURBING MEMBRANE FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
CC SHOWS ANTICANCER ACTIVITY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC Amidation; Antibiotic.
FW MOD RES 17
SQ SEQUENCE 17 AA; 1769 MW; F6AC0A580428EAFD CRC64;

Query Match 20.4%; Score 20; DB 1; Length 17;
Best Local Similarity 50.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 10 ATHYAESV 17
Db 10 AGHIASSI 17

RESULT 10
ATPB_CANFA
ID ATPB_CANFA STANDARD; PRT; 19 AA.
AC P99504;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase beta chain, mitochondrial (EC 3.6.3.14) (Fragment).
GN ATP5B.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-APPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
```

```
RL Electrophoresis 18:2795-2802(1997).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC
CC SUBUNIT.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR HSC-2DPAGE; P99504; DOG.
DR InterPro; IPR000194; ATPase a/bcentre.
DR PROSITE; PS00152; ATPASE ALPHA BETA; PARTIAL.
KW ATP synthesis; CF(1); Hydrogen ion transport; Hydrolase; ATP-binding;
KW Mitochondrion.
FT UNSURE 8
FT NON TER 19
FT UNSURE 17
FT NON TER 19
SQ SEQUENCE 19 AA; 1871 MW; BB9C163FDC60BB42 CRC64;

Query Match 20.4%; Score 20; DB 1; Length 19;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 10 ATHYAESVKG 19
Db 1 ATQTSPPKRG 10

RESULT 11
LSK2_LEUMA
ID LSK2_LEUMA STANDARD; PRT; 10 AA.
AC P09039;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Leucosulfakinin-II (LSK-II).
OS Leucophaea maderae (Madeira cockroach), and
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988, 6978;
RN [1]
SEQUENCE.
RC SPECIES=L.maderae; PubMed=3778455;
RX MEDLINE=87048769; PubMed=3778455;
RA Nachman R.J., Holman G.M., Cook B.J., Haddon W.F., Ling N.;
RT "Leucosulfakinin-II, a blocked sulfated insect neuropeptide with
RT homology to cholecystokinin and gastrin."
RL Biochem. Biophys. Res. Commun. 140:357-364(1986).
RN [2]
SEQUENCE.
RC SPECIES=P.americana; TISSUE=Corpora cardiaca;
RX MEDLINE=90137190; PubMed=2615921;
RA Veenstra J.A.;
RT "Isolation and structure of two gastrin/CCK-like neuropeptides from
RT the American cockroach homologous to the leucosulfakinins."
RL Neuropeptides 14:145-149(1989).
CC -1- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF
CC THE COCKROACH HINGUT. STIMULATES MUSCLE CONTRACTION OF HINDGUT.
CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR PIR; A26335; GMR0L2.
DR PIR; B60656; B60656.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Hormone; Amidation; Sulfation.
FT MOD RES 1
FT MOD RES 5
FT MOD RES 5
FT MOD RES 10
FT MOD RES 10
```

SQ SEQUENCE 10 AA; 1255 MW; 9B4F5391E86B5AAA CRC64;  
 Query Match 19.4%; Score 19; DB 1; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 KSDNY 9  
 Db 1 QSDDY 5  
 RESULT 12  
 TKNB\_RANCA  
 ID TKNB\_RANCA STANDARD; PRT; 10 AA.  
 AC P22689;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Ranachykinin B (RTK B).  
 OS Rana catesbeiana (Bull frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.  
 OX NCBI\_TaxID=8400;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Brain;  
 RX MEDLINE=91254337; PubMed=2043143;  
 RA Kozawa H., Hino J., Minamino N., Kangawa K., Matsuo H.;  
 RT "Isolation of four novel tachykinins from frog (Rana catesbeiana)  
 brain and intestine.";  
 RL Biochem. Biophys. Res. Commun. 177:588-595 (1991).  
 RN [2]  
 RP SEQUENCE.  
 RC TISSUE=Intestine;  
 RX MEDLINE=94023216; PubMed=8210506;  
 RA Kangawa K., Kozawa H., Hino J., Minamino N., Matsuo H.;  
 RT "Four novel tachykinins in frog (Rana catesbeiana) brain and  
 intestine.";  
 RL Regul. Pept. 46:81-88 (1993).  
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 MUSCLES.  
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR; JE0427; JE0427.  
 DR PIR; B61033; B61033.  
 DR InterPro; IPR002040; Tachykinin.  
 DR PROSITE; PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation.  
 FT MOD RES 10 10  
 FT SEQUENCE 10 AA; 1210 MW; 917E556B59D5BAB5 CRC64;  
 Query Match 19.4%; Score 19; DB 1; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 KSDNY 9  
 Db 2 KSDSF 6  
 RESULT 13  
 UN37\_CLOPA  
 ID UN37\_CLOPA STANDARD; PRT; 14 AA.  
 AC P81358;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Unknown protein CP 37 from 2D-page (Fragment).  
 OS Clostridium pasteurianum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1501;

RN [1]  
 RP SEQUENCE.  
 RC STRAIN=W5;  
 RX MEDLINE=98291870; PubMed=9629918;  
 RA Flengsrud R., Skjeldal L.;  
 RT "Two-dimensional gel electrophoresis separation and N-terminal  
 sequence analysis of proteins from Clostridium pasteurianum W5.";  
 RL Electrophoresis 19:802-806 (1998).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 PROTEIN IS: 5.8, ITS MW IS: 44.7 kDa.  
 FT NON TER 14 14  
 FT SEQUENCE 14 AA; 1579 MW; 05384662DEF89210 CRC64;  
 Query Match 19.4%; Score 19; DB 1; Length 14;  
 Best Local Similarity 41.7%; Pred. No. 2.5e+03;  
 Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 Qy 8 NYATHYAESVKG 19  
 Db 3 NKNXIEDQGVKG 14  
 RESULT 14  
 MILT\_ONCKE  
 ID MILT\_ONCKE STANDARD; PRT; 15 AA.  
 AC P81037;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Miltpain (EC 3.4.22.-) (Fragment).  
 OS Oncorhynchus keta (Chum salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Proacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8018;  
 RN [1]  
 RP SEQUENCE, AND CHARACTERIZATION.  
 RX MEDLINE=97397031; PubMed=9253183;  
 RA Kawabata C., Ichishima E.;  
 RT "Miltpain, new cysteine proteinase from the milt of chum salmon,  
 Oncorhynchus keta.";  
 RL Comp. Biochem. Physiol. 117B:445-452 (1997).  
 CC -1- FUNCTION: CYSTEINE PROTEINASE THAT HYDROLYSES BASIC PROTEINS.  
 CC -1- HYDROLYZE BASIC PROTEINS SUCH AS HISTONE, SALMINE AND CLUPAINE BUT.  
 CC NOT MILK CASEIN.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE WITH BASIC RESIDUES AT  
 P2 AND P1.  
 CC MEROPS; C01.093; --  
 DR Hydrolase.  
 FT NON TER 15 15  
 FT SEQUENCE 15 AA; 1730 MW; 766B7771C0F898E7 CRC64;  
 Query Match 19.4%; Score 19; DB 1; Length 15;  
 Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 13 YAESV 17  
 Db 6 YAEV 10  
 RESULT 15  
 NIS1\_ANAVA  
 ID NIS1\_ANAVA STANDARD; PRT; 15 AA.  
 AC Q44507;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cysteine desulfurase 1 (EC 4.4.1.-) (Nitrogenase metalloclusters  
 biosynthesis protein nifS) (Fragment).  
 GN NIFS1 OR NIFS.  
 OS Anabaena variabilis.  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.

OX NCBI\_TaxID=1172;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PCC 7937 / ATCC 29413;  
RA Monnerjahn U., Boehme H.;  
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP EXPRESSION PATTERN.  
RC STRAIN=PCC 7937 / ATCC 29413;  
RX MEDLINE=96016168; PubMed=7568132;  
RA Thiel T., Lyons E.M., Erker J.C., Ernst A.;  
RT "A second nitrogenase in vegetative cells of a heterocyst-forming  
cyanobacterium."  
RL Proc. Natl. Acad. Sci. U.S.A. 92:9358-9362(1995).  
CC -!- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR ATOMS FROM  
CYSTEINE TO PRODUCE ALANINE. SEEMS TO PARTICIPATE IN THE  
BIOSYNTHESIS OF THE NITROGENASE METALLOCLUSTERS BY PROVIDING THE  
INORGANIC SULFUR REQUIRED FOR THE FE-S CORE FORMATION (BY  
SIMILARITY).  
CC -!- COPACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- MISCELLANEOUS: BELONGS TO THE NIF1 GENE CLUSTER WHICH IS EXPRESSED  
IN HETEROCYSTS UNDER ANAEROBIC AND AEROBIC CONDITIONS.  
CC -!- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
AMINOTRANSFERASES. NIFS/ISCS SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X69898; CAA49523.1; --  
DR InterPro; IPR000192; Aminotransfv.  
DR PROSITE; PS00595; AA\_TRANSFER\_CLASS\_5; PARTIAL.  
KW Nitrogen fixation; Lyase; Pyridoxal phosphate.  
FT NON TER 15  
SQ SEQUENCE 15 AA; 1684 MW; 08B8F106DE65547D CRC64;  
  
Query Match 19.4%; Score 19; DB 1; Length 15;  
Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 7 DNYAT 11  
Db 7 DNNAT 11

Search completed: April 22, 2003, 13:48:59  
Job time : 27 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:45:28 ; Search time 28 Seconds  
(without alignments)  
139.818 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98  
Sequence: 1 EIRLKSNDYATHAESVKG 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 5270

Minimum DB seq length: 0  
Maximum DB seq length: 19

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: sp\_archea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	32.7	19	15	Q73129 human immun
2	30	30.6	15	4	P83219 homo sapien
3	29	29.6	15	15	Q91BM5 human immun
4	28	28.6	13	4	Q9UC27 homo sapien
5	27	27.6	19	2	Q9L434 salmonella
6	26	26.5	13	2	Q9KHJ4 caulobacter
7	26	26.5	15	1	Q9UWH6 thermococu
8	26	26.5	15	6	Q9TRT6 sus scrofa
9	26	26.5	15	6	Q9TRT4 sus scrofa
10	26	26.5	19	12	O84863 unidentified
11	25	25.5	17	1	Q9V2X3 methanobact
12	24	24.5	13	5	Q9TRW4 tityus serr
13	24	24.5	14	4	Q93057 homo sapien
14	24	24.5	14	11	P70319 mus muscucu
15	23	23.5	9	2	Q93LE4 heliobacill
16	23	23.5	10	2	P96423 pseudomonas

17	23	23.5	14	2	P96347	P96347 helicobacte
18	23	23.5	15	15	Q8UP12	Q8UP12 human immun
19	23	23.5	17	2	Q9R4H9	Q9R4H9 bordetella
20	23	23.5	18	1	Q9UWJ7	Q9UWJ7 archaeoglob
21	23	23.5	18	13	Q9PRX1	Q9PRX1 struthio ca
22	22	22.4	16	2	Q9R4J4	Q9R4J4 pseudomonas
23	22	22.4	17	2	Q9R4P9	Q9R4P9 pseudomonas
24	22	22.4	17	5	Q9V973	Q9V973 drosophila
25	22	22.4	18	2	Q9RSF6	Q9RSF6 helicobacte
26	22	22.4	18	6	Q9TRF1	Q9TRF1 bos taurus
27	22	22.4	18	8	Q9TRF9	Q9TRF9 theileria a
28	21	21.4	11	2	O56972	O56972 versinia pe
29	21	21.4	11	4	O60842	O60842 homo sapien
30	21	21.4	15	1	Q9UWH4	Q9UWH4 pyrococcus.
31	21	21.4	15	5	P83246	P83246 asaphis dic
32	21	21.4	15	6	Q9TRT3	Q9TRT3 sus scrofa
33	21	21.4	16	2	O53399	O53399 bacillus th
34	21	21.4	16	6	Q9TR18	Q9TR18 oryctolagus
35	21	21.4	16	11	Q9CTB2	Q9CTB2 mus muscucu
36	21	21.4	17	2	Q52210	Q52210 staphylococ
37	21	21.4	18	2	Q47337	Q47337 escherichia
38	21	21.4	18	7	Q9TNP1	Q9TNP1 mus sp. mhc
39	21	21.4	18	11	Q9QUX0	Q9QUX0 mus sp. tum
40	21	21.4	19	2	O53545	O53545 shigella so
41	21	21.4	19	6	Q9TRF7	Q9TRF7 oryctolagus
42	21	21.4	19	10	Q43370	Q43370 allium cepa
43	20	20.4	10	2	P96305	P96305 alteromonas
44	20	20.4	14	2	O54394	O54394 streptomyce
45	20	20.4	16	9	Q38407	Q38407 bacterioph

## ALIGNMENTS

RESULT 1

Q73129 PRELIMINARY; PRT; 19 AA.  
AC Q73129;  
DT 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SAMPLE P4.25-15;  
RA Wellinsky S.M., Korber B.T.M., Neumann A.U., Daniels M., Kunstman K.J.,  
RA Whetzel A.J., Cao Y., Ho D.D., Safritz J.T., Koup R.A.;  
RT "Adaptive Evolution of Human Immunodeficiency Virus Type 1 During the  
RT Natural Course of Infection";  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SAMPLE P4.25-15;  
RA McDonald D.;  
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U36027; AAA97722.1; -;  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
SQ SEQUENCE 19 AA; 2066 MW; E117102D52800BF4 CRC64;  
Query Match 32.7%; Score 32; DB 15; Length 19;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EIRLKSNDY 10  
|: :|||:  
Db 7 EVVIRSDNFA 16

```
RESULT 2
P83219 ID P83219 PRELIMINARY; PRT; 15 AA.
AC P83219
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Calcitermin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=NASAL MUCUS;
RX MEDLINE=21413725; PubMed=11522286;
RA Cole A.M., Kim Y.-H., Tahk S., Hong T., Weis P., Waring A.J., Ganz T.;
RT "Calcitermin, a novel antimicrobial peptide isolated from human airway
secretions.";
RL FEBS Lett. 504:5-10(2001).
CC -1- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY AGAINST C.ALBICANS AND IS
CC ALSO ACTIVE AGAINST E.COLI AND P.AERUGINOSA BUT NOT
CC L.MONOCYTOGENES AND S.AUREUS.
CC -1- MASS SPECTROMETRY: MW=1689.9; METHOD=MALDI.
CC -1- MISCELLANEOUS: OPTIMUM PH IS 5.4.
CC -1- MISCELLANEOUS: ANTIMICROBIAL ACTIVITY IS ENHANCED BY ZNCL2.
CC -1- SIMILARITY: TO THE C-TERMINUS OF HUMAN CALGRANULIN C.
KW Antibiotic; Fungicide; Zinc; Metal-binding.
FT DOMAIN 9 13 ZINC-BINDING (POTENTIAL).
SQ SEQUENCE 15 AA; 1689 MW; CE4EC8DDFE861663 CRC64;

Query Match 30.6%; Score 30; DB 4; Length 15;
Best Local Similarity 54.5%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 IRLKSDNYATH 12
DB 3 IALKAARVYHTH 13

RESULT 3
Q91BMS ID Q91BMS PRELIMINARY; PRT; 15 AA.
AC Q91BMS
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Envelope glycoprotein, V3 region (Fragment).
GN ENV
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=111676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95111504; PubMed=7815476;
RA Ahmad N., Baroudy B.M., Baker R.C., Chappey C.;
RT "Genetic analysis of human immunodeficiency virus type 1 envelope V3
RT region isolates from mothers and infants after perinatal
RT transmission.";
RL J. Virol. 69:1001-1012(1995).
DR EMBL; U16595; AAA64161.1; -.
FT NON_TER 1 1
SQ SEQUENCE 15 AA; 1794 MW; 747CGA62800EF4A1 CRC64;

Query Match 29.6%; Score 29; DB 15; Length 15;
Best Local Similarity 40.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSDNYA 10
DB 4 EVIIRSDNFS 13

RESULT 4
Q9UC27 ID Q9UC27 PRELIMINARY; PRT; 13 AA.
AC Q9UC27
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE 94 kDa epididymal cyokeratin-like protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=95226590; PubMed=7536047;
RA Boue F., Duquenne C., Lassealle B., Lefevre A., Finaz C.;
RT "FLB1, a human protein of epididymal origin that is involved in the
RT sperm-oocyte recognition process.";
RL Biol. Reprod. 52:267-278(1995).
SQ SEQUENCE 13 AA; 1610 MW; 000A693B81B7A45B CRC64;

Query Match 28.6%; Score 28; DB 4; Length 13;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 NYATHY 13
DB 1 NYSTY 6

RESULT 5
Q9L434 ID Q9L434 PRELIMINARY; PRT; 19 AA.
AC Q9L434
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 2.4 kDa protein (Fragment).
GN PUTATIVE YCIE.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CS2;
RX MEDLINE=21160188; PubMed=11260470;
RA Robbe-Saule V., Coynault C., Ibanez-Ruiz M., Hermant D., Norel F.;
RT "Identification of a non-haem catalase in Salmonella and its
RT regulation by RpoS (sigma S).";
RL Mol. Microbiol. 39:1533-1545(2001).
DR EMBL; AJ251362; CAB71038.1; -.
KW Hypothetical protein.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2448 MW; 05D3DB83C7CBE7E5 CRC64;

Query Match 27.6%; Score 27; DB 2; Length 19;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 NYATHY 13
DB 2 NYTEHY 7

RESULT 6
Q9KHJ4 ID Q9KHJ4 PRELIMINARY; PRT; 13 AA.
AC Q9KHJ4
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Proteolysis tag (Fragment).
```



OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=10884408;  
 RX MEDLINE=20345063; Shapiro L., Williams K.P.;  
 RA Keller K.C., Shapiro L., Williams K.P.;  
 RT "tRNAs that encode proteolysis-inducing tags are found in all known  
 bacterial genomes: A two-piece tRNA functions in Caulobacter.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:7778-7783(2000).  
 DR EMBL: AF255738; AAF87998.1; -.  
 FT NON TER 1  
 SQ SEQUENCE 13 AA; 1368 MW; CE5F60C57FCE1B1D CRC64;  
 Query Match 26.5%; Score 26; DB 2; Length 13;  
 Best Local Similarity 44.4%; Pred. No. 9.4e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy 6 SDNYATHYA 14  
 :|||:  
 Db 2 NDNFAEFA 10  
 RESULT 7  
 Q9UWH6 PRELIMINARY; PRT; 15 AA.  
 ID Q9UWH6  
 AC Q9UWH6  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE keto-valine-ferredoxin oxidoreductase alpha-subunit, VOR alpha  
 (Fragment).  
 OS Thermococcus litoralis.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Thermococcus.  
 OX NCBI\_TaxID=2265;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96146528; PubMed=8550513;  
 RA Heider J., Mai X., Adams M.W.;  
 RT "Characterization of 2-ketoisovalerate ferredoxin oxidoreductase, a  
 new and reversible coenzyme A-dependent enzyme involved in peptide  
 fermentation by hyperthermophilic archaea.";  
 RL J. Bacteriol. 178:780-787(1996).  
 SQ SEQUENCE 15 AA; 1549 MW; OC1766910B222944 CRC64;  
 Query Match 26.5%; Score 26; DB 1; Length 15;  
 Best Local Similarity 45.5%; Pred. No. 1.1e+03;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 Qy 3 RLKSDNYATHY 13  
 :|||:  
 Db 3 KVSNGYAAAY 13  
 RESULT 8  
 Q9TRT6 PRELIMINARY; PRT; 15 AA.  
 ID Q9TRT6  
 AC Q9TRT6  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE Pyruvate dehydrogenase (EC 1.2.4.1) (Fragment).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92144618; PubMed=1737046;  
 RA Koike K., Urata Y., Goto S.;  
 RT "Proteinase-catalyzed activation of porcine heart muscle pyruvate

RT dehydrogenase and identification of its cleavage site.";  
 RL Biochim. Biophys. Acta 1118:223-230(1992).  
 SQ SEQUENCE 15 AA; 1787 MW; DB1265CDF7737E16 CRC64;  
 Query Match 26.5%; Score 26; DB 6; Length 15;  
 Best Local Similarity 71.4%; Pred. No. 1.1e+03;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 EIRLKSD 7  
 :|||:  
 Db 7 EVRSKSD 13  
 RESULT 9  
 Q9TRT4 PRELIMINARY; PRT; 15 AA.  
 ID Q9TRT4  
 AC Q9TRT4  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE Pyruvate dehydrogenase (EC 1.2.4.1) (Fragment).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92144618; PubMed=1737046;  
 RA Koike K., Urata Y., Goto S.;  
 RT "Proteinase-catalyzed activation of porcine heart muscle pyruvate  
 dehydrogenase and identification of its cleavage site.";  
 RL Biochim. Biophys. Acta 1118:223-230(1992).  
 SQ SEQUENCE 15 AA; 1757 MW; 94D6BFE4B126598 CRC64;  
 Query Match 26.5%; Score 26; DB 6; Length 15;  
 Best Local Similarity 71.4%; Pred. No. 1.1e+03;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 EIRLKSD 7  
 :|||:  
 Db 3 EVRSKSD 9  
 RESULT 10  
 ID Q84863 PRELIMINARY; PRT; 19 AA.  
 AC Q84863  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical 2.1 kDa protein (Fragment).  
 OS unidentified human poliovirus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OX NCBI\_TaxID=40278;  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=3031313;  
 RX MEDLINE=87169734; Kuge S., Saito I., Nomoto A.;  
 RA Kuge S., Saito I., Nomoto A.;  
 RT "Primary structure of poliovirus defective-interfering particle  
 genomes and possible generation mechanisms of the particles.";  
 RL J. Mol. Biol. 192:473-487(1986).  
 DR EMBL: M30221; AAA66829.1; -.  
 KW Hypothetical protein.  
 FT NON TER 1  
 FT NON TER 19  
 SQ SEQUENCE 19 AA; 2106 MW; F774197992500AC8 CRC64;  
 Query Match 26.5%; Score 26; DB 12; Length 19;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 8 NVATH 12.

```

Db          |||||
           9 NYTH 13

RESULT 11
Q9V2X3      PRELIMINARY;      PRT;      17 AA.
AC Q9V2X3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE RNA polymerase subunit B', (Fragment).
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88198019; PubMed=2834336;
RA Schallenberg J., Moss M., Truss M., Reiser W., Thomm M., Stetter K.O.,
RA Klein A.;
RT "Cloning and physical mapping of RNA polymerase genes from
RT Methanobacterium thermoautotrophicum and comparison of homologies and
RT gene orders with those of RNA polymerase genes from other Methanogenic
RT archaeobacteria."
RL J. Bacteriol. 170:2247-2253(1988).
DR EMBL; M20391; AAA72654.1; -.
FT NON TER 17
SQ SEQUENCE 17 AA; 2137 MW; 0C03CD9B6D780560 CRC64;

Query Match      25.5%; Score 25; DB 1; Length 17;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 11 THYAES 16
   |||||
Db 4 THYPEN 9

RESULT 12
Q9TWR4      PRELIMINARY;      PRT;      13 AA.
AC Q9TWR4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Peptide T-BRADYKININ potentiator.
OS Tityus serrulatus (Brazilian scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Euthoidea; Buthidae; Tityus.
OX NCBI_TaxID=6887;
RN [1]
RP SEQUENCE.
RX MEDLINE=94024945; PubMed=8212046;
RA Ferreira L.A., Alves E.W., Henriques O.B.;
RA "Peptide T, a novel bradykinin potentiator isolated from Tityus
RA serrulatus scorpion venom."
RL Toxicon 31:941-947(1993).
SQ SEQUENCE 13 AA; 1604 MW; 35770B0644FC02D7 CRC64;

Query Match      24.5%; Score 24; DB 5; Length 13;
Best Local Similarity 44.4%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 KSDNYATHY 13
   |||||
Db 1 KKGDPVEY 9

RESULT 13
Q93057      PRELIMINARY;      PRT;      14 AA.
ID Q93057
AC Q93057;

```

```

DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE B cell specific activator protein BSAP (PAX5) (Fragment).
OS PAX5 OR PAX-5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Iida S., Rao P.H., Nallasiwam P., Hibshoosh H., Butler M., Louie D.C.,
RA Dyonin V., Ohno H., Chaganti R.S.K., Dalla-Favera R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96234102; PubMed=8650231;
RA Busslinger M., Klix N., Pfeiffer P., Graninger P.G., Kozmik Z.;
RT "Deregulation of PAX-5 by translocation of the Emu enhancer of the Igh
RT locus adjacent to two alternative PAX-5 promoters in a diffuse large-
RT cell lymphoma."
RL Proc. Natl. Acad. Sci. U.S.A. 93:6129-6134(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=MARGINAL ZONE LYMPHOMA;
RA Morrison A.M., Jaeger U., Chott A., Haas O.A., Schebesta M.,
RA Busslinger M.;
RT "Deregulated PAX-5 transcription from a translocated Igh promoter in
RT marginal zone lymphoma."
RL Blood 0:0-0(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=21354098; PubMed=11460166;
RA Pasqualucci L., Neumeister P., Goossens T., Nanjangud G.,
RA Chaganti R.S.K., Kuppers R., Dalla-Favera R.;
RT "Hypermutation of multiple proto-oncogenes in B-cell diffuse large-
RT cell lymphomas."
RL Nature 412:341-346(2001).
DR EMBL; U62539; AAB09533.1; -.
DR EMBL; U56836; AAB16833.1; -.
DR EMBL; AF074913; AAC69188.1; -.
DR EMBL; AF386791; AAK70870.1; -.
FT NON TER 14
SQ SEQUENCE 14 AA; 1683 MW; E113518FA19953A1 CRC64;

Query Match      24.5%; Score 24; DB 4; Length 14;
Best Local Similarity 45.5%; Pred. No. 2.2e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EIRLKSQNYAT 11
   |||||
Db 2 EIRKHDPPAS 12

RESULT 14
P70319      PRELIMINARY;      PRT;      14 AA.
ID P70319;
AC P70319;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE B-cell specific transcription factor (Fragment).
GN PAX-5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=129/SV;
RT "Deregulation of PAX-5 by translocation of the Emu enhancer of the Igh

```

RT locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:6129-6134(1996).  
 DR EMBL; U56838; AAB16835.1; -.  
 FT NON TER 14 14  
 SQ SEQUENCE 14 AA; 1683 MW; E113518PA19953A1 CRC64;

Query Match 24.5%; Score 24; DB 11; Length 14;  
 Best Local Similarity 45.5%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EIRLKSDNYAT 11  
 ||| |  
 Db 2 EIHCKHDPFAS 12

# RESULT 15

ID Q93LE4 PRELIMINARY; PRT; 9 AA.  
 AC Q93LE4;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Proteolysis tag (Fragment).  
 OS Helicobacter mobilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Helicobacteriaceae; Helicobacter.  
 OX NCBI\_TaxID=28064;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20063247; PubMed=10592213;  
 RA Williams K.P.;  
 RT "The tRNA website."  
 RL Nucleic Acids Res. 28:168-168(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Williams K.P.;  
 RT "Phylogenetic analysis of tRNA."  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY040838; AAK83526.1; -.  
 FT NON TER 1 1  
 SQ SEQUENCE 9 AA; 937 MW; 7CD8D72DCB544AAB CRC64;

Query Match 23.5%; Score 23; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 6.7e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 DNYA 10  
 ||||  
 Db 3 DNYA 6

Search completed: April 22, 2003, 13:49:34  
 Job time : 30 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:42:32 ; Search time 73 Seconds  
(without alignments)  
34.682 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98

Sequence: 1 EIRKSDNYTHAIVSVKG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 278369

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq 101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	19	21	AAV32258
2	91	92.9	19	22	AAV32258
3	80	81.6	17	21	AAV30033
4	74	75.5	16	20	AAV40694
5	70	71.4	19	21	AAV10002
6	68	69.4	19	23	AAV84529
7	65	66.3	19	23	AAV14988
8	64	65.3	19	20	AAV94737
9	52	53.1	17	23	AAV82539
10	49	50.0	19	16	AAV74934

11	49	50.0	19	23	AAU70371	Human heavy chain
12	48	49.0	19	23	AAU70380	Mouse heavy chain
13	47	48.0	19	16	AAV74933	H-CDR-2 of anti-id
14	46	46.9	17	21	AAV30034	Scaffold protein S
15	45	45.9	19	18	AAW23392	Mutant 3 of the CD
16	43	43.9	17	16	AAV75491	Mouse antibody var
17	43	43.9	17	21	AAV39862	Anti-hil12 antibody
18	43	43.9	17	23	AAU82463	Llama CDR2 region
19	43	43.9	18	21	AAV52188	Human anti-HB8 ant
20	42	42.9	16	23	AAU82611	Llama CDR2 region
21	42	42.9	17	20	AAW90323	Human anti-idiotyp
22	42	42.9	17	22	AAE10817	Human antibody CAT
23	42	42.9	17	23	AAU82537	Llama CDR2 region
24	42	42.9	17	23	AAU82550	Llama CDR2 region
25	41	41.8	17	14	AAV32117	Heavy chain CDR2 r
26	41	41.8	17	20	AAW90330	Human anti-idiotyp
27	41	41.8	17	21	AAV39846	Anti-hil12 antibod
28	41	41.8	17	22	AAV67496	Human heavy chain
29	41	41.8	17	23	AAU82488	Llama CDR2 region
30	41	41.8	17	23	AAU82543	Llama CDR2 region
31	41	41.8	19	18	AAV33394	Mutant 5 of the CD
32	41	41.8	19	21	AAV5202	Anti-platelet glyco
33	41	41.8	19	23	AAU9858	Mouse Hefi antibod
34	40	40.8	16	20	AAV40695	A3 derivative #16,
35	40	40.8	16	23	AAU82466	Llama CDR2 region
36	40	40.8	17	18	AAV16651	Anti-cancer specif
37	40	40.8	17	20	AAV05054	Tumour antigen ant
38	40	40.8	17	21	AAV39819	Anti-hil12 antibod
39	40	40.8	17	21	AAV39833	Anti-hil12 antibod
40	40	40.8	17	21	AAV39868	Anti-hil12 antibod
41	40	40.8	17	22	AAV95213	Anti-platelet glyco
42	40	40.8	17	22	AAV99156	Human scFv clone 2
43	40	40.8	17	22	AAV84980	D12 scFv CDR H2 re
44	40	40.8	17	22	AAV61294	Anti-TANGO 268 scF
45	40	40.8	17	22	AAV65311	Anti-IL-18 antibod

## ALIGNMENTS

RESULT 1  
AAV32258  
ID AAV32258 standard, Peptide; 19 AA.  
XX  
AC AAV32258;  
XX  
DT 15-FEB-2000 (first entry)  
XX  
DE Light chain CDR H2 of mouse anti-CD23 MAb C11.  
XX  
KW CD23, FCER11; IgE receptor; monoclonal antibody; C11; mouse;  
KW monoclonal antibody; chimeric antibody; humanised antibody;  
KW complementarity determining region; CDR; autoimmune disease;  
KW inflammation; arthritis; lupus erythematosus; multiple sclerosis;  
KW Hashimoto's thyroiditis; diabetes; uveitis; dermatitis; psoriasis;  
KW urticaria; nephrotic syndrome; glomerulonephritis;  
KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;  
KW Sjogren's syndrome; allergy; asthma; rhinitis; eczema; insulinitis;  
KW graft-versus-host disease; COPD; bronchitis; diabetes;  
KW B-cell malignancy; therapy.  
XX  
OS Mus musculus.

PN WO9558679-A1.  
XX.  
PD 18-NOV-1999.  
XX  
PF 07-MAY-1999; 99WO-GB01434.  
XX  
PR 09-MAY-1998; 98GB-0009839.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX

PI Bonney J, Crowe SJ, Ellis JH, Rapson NT, Shearin J;  
 XX WPI; 2000-053101/04.  
 DR N-PSDB; AAZ34743.  
 XX  
 PT Cell receptor specific antibodies useful for treating e.g. arthritis,  
 PT diabetes, multiple sclerosis and psoriasis -  
 XX Claim 1; Page 40; 81pp; English.  
 XX  
 CC This sequence represents complementarity determining region 2  
 CC (CDR H2) of the heavy chain of murine anti-CD23 (FCER1I) monoclonal  
 CC antibody C11 (see also AAY2263). The invention provides altered  
 CC antibodies, such as chimeric or humanised antibodies, which comprise  
 CC sufficient of the amino acid sequences of C11 light and heavy chain  
 CC CDRs (see AAY2254-59) to render them capable of binding to the CD23  
 CC type II molecule expressed on haematopoietic cells. The antibodies  
 CC are used to block soluble CD23 formation for treatment of arthritis,  
 CC lupus erythematosus, Hashimoto's thyroiditis, multiple sclerosis,  
 CC diabetes, uveitis, dermatitis, psoriasis, urticaria, nephrotic  
 CC syndrome, glomerulonephritis, inflammatory bowel disease, ulcerative  
 CC colitis, Crohn's disease, Sjogren's syndrome, allergies, allergic  
 CC asthma, intrinsic asthma, acute asthmatic exacerbation, rhinitis,  
 CC eczema, graft-versus-host disease, COPD, insulinitis, bronchitis  
 CC (particularly chronic bronchitis) or diabetes (particularly type 1  
 CC diabetes), and B-cell malignancies (claimed). They are also useful  
 CC for studying interactions between CD23 and various ligands and  
 CC determining the binding agents.  
 XX Sequence 19 AA;  
 XX  
 Query Match 100.0%; Score 98; DB 21; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLSDNYATHYAESVKG 19  
 Db 1 EIRLSDNYATHYAESVKG 19  
 XX  
 RESULT 2  
 AAB35297  
 ID AAB35297 standard; Peptide; 19 AA.  
 XX  
 AC AAB35297;  
 XX  
 DT 08-MAY-2001 (first entry)  
 XX  
 DE Murine PSCA antibody fragment #5.  
 XX  
 KW Prostate stem cell antigen; PSCA; human; mouse; prostate cancer;  
 KW diagnosis; treatment; chromosome 8q24.2.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200105427-A1.  
 XX  
 PD 25-JAN-2001.  
 XX  
 PF 20-JUL-2000; 2000WO-US19967.  
 XX  
 PR 20-JUL-1999; 99US-0359326.  
 PR 03-MAY-2000; 2000US-0564329.  
 XX  
 XX (REG ) UNIV CALIFORNIA.  
 PA (UROC ) UROGENESYS.  
 XX  
 PI Reiter R, Witte O, Saffran DC, Jakobovits A;  
 XX WPI; 2001-159478/16.  
 XX  
 PT Antibodies binding to prostate stem cell antigen inhibit the growth of  
 PT cancer cells and are used to detect and treat prostate, pancreatic or

PT bladder cancers -  
 XX Example 21; Fig 61; 229pp; English.  
 XX  
 CC The present invention describes a method of treating cancer associated  
 CC with prostate stem cell antigen (PSCA) by administering an antibody which  
 CC selectively binds to PSCA and inhibits the growth of the cancer cells.  
 CC The PSCA gene is found on human chromosome 8q24.2. The invention provides  
 CC the human and murine PSCA protein and coding sequences, which can be used  
 CC not only in the treatment of, but also in detection and prognosis of  
 CC prostate cancer.  
 XX Sequence 19 AA;  
 XX  
 Query Match 92.9%; Score 91; DB 22; Length 19;  
 Best Local Similarity 89.5%; Pred. No. 2e-08;  
 Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EIRLSDNYATHYAESVKG 19  
 Db 1 EIRLSDNYATHYAESVKG 19  
 XX  
 RESULT 3  
 AAB30033  
 ID AAB30033 standard; Peptide; 17 AA.  
 XX  
 AC AAB30033;  
 XX  
 DT 09-FEB-2001 (first entry)  
 XX  
 DE Scaffold protein SCA A3 peptide SEQ ID NO: 94.  
 XX  
 KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;  
 KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;  
 KW diabetic retinopathy; atherosclerosis.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2000060070-A1.  
 XX  
 PD 12-OCT-2000.  
 XX  
 PF 01-APR-1999; 99WO-EP02283.  
 XX  
 PR 01-APR-1999; 99WO-EP02283.  
 XX  
 PA (INNO-) INNOGENETICS NV.  
 XX  
 PI Desmet J, Hufton S, Hoogenboom H, Sablon E;  
 XX WPI; 2000-665002/64.  
 XX  
 PT Scaffold composed of single-chain polypeptide having beta sandwich  
 PT architecture carrying new and randomized peptide sequences useful as  
 PT supporting framework and carrying antigen- or receptor binding  
 PT fragments -  
 XX Disclosure; Page 14; 68pp; English.  
 XX  
 CC The present invention is concerned with producing scaffold proteins  
 CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be  
 CC used as a scaffold to bind antigen- or receptor-binding fragments. These  
 CC can be used in the treatment of diseases such as cancer,  
 CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and  
 CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the  
 CC production of the proteins of the invention.  
 XX Sequence 17 AA;  
 XX  
 Query Match 81.6%; Score 80; DB 21; Length 17;  
 Best Local Similarity 88.2%; Pred. No. 1.2e-06;  
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVK 19  
 |||:|||||  
 Db 1 RLNSDNFATHYAESVK 17

RESULT 4  
 AAY40694  
 ID AAY40694 standard; peptide; 16 AA.  
 AC AAY40694;

XX DT 01-DEC-1999 (first entry)

XX A3 derivative #15, beta strand of scaffold protein structure.

XX Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;  
 KW tumour; chemotherapeutic agent.

XX Synthetic.

XX EP947582-A1.

XX 06-OCT-1999.

XX 31-MAR-1998; 98EP-0870065.

XX 31-MAR-1998; 98EP-0870065.

XX (INNO-) INNOGENETICS NV.

XX Desmet J, Hufton S, Hoogenboom H, Sablon E;

XX WPI; 1999-542958/46.

XX New scaffold protein, useful for stabilizing antigens used as vaccines

XX Disclosure; Page 6; 105pp; English.

XX Sequences AAY40680-Y40703 are functionally equivalent derivatives of the  
 CC A3 peptide (AAY40605) which forms part of a scaffold protein. A3 is a  
 CC beta strand peptide which forms part of a beta sheet. Peptides  
 CC (AAY40601-Y40609) together form a single-chain scaffold protein which  
 CC contains at least 1 disulfide bond, contains less than 10% alpha helix  
 CC and contains at least 6 beta-strands. The scaffold protein is constructed  
 CC of beta strands S1-S6, and may also include beta strands A1-A3, or any  
 CC functionally equivalent derivative of these sequences. The beta strands  
 CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to  
 CC the next by hydrogen bonds, which generate a beta sandwich architecture.  
 CC If the additional beta strands A1-A3 are included in the structure the  
 CC scaffold is constructed of two beta sheets, with the structures  
 CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each  
 CC other via amino acid loops, where at least one of the loops binds to a  
 CC receptor or antigen. The scaffold protein is used to stabilize antigens  
 CC or whole proteins such as receptors, or their fragments. It may be used  
 CC to bind two separate molecules. For example, one surface of the scaffold  
 CC may be bound to a protein which binds to a tumour antigen. This will  
 CC target the complex to tumour cells. Another surface may be bound to a  
 CC cytotoxic molecule or an autoimmune antibody which may then kill the  
 CC tumour cells. Therefore the scaffold protein may be used to target  
 CC chemotherapeutic agents to specific cells. It may also be used to  
 CC stabilize individual peptides in a peptide library and may be used in  
 CC diagnostic techniques, and to stabilize antigens used as vaccines.

XX Sequence 16 AA;

Query Match 75.5%; Score 74; DB 20; Length 16;  
 Best Local Similarity 87.5%; Pred. No. 1.1e-05;  
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVK 18  
 |||:|||||

Db 1 RLNSDNFATHYAESVK 16

RESULT 5

AAB10002  
 ID AAB10002 standard; Protein; 19 AA.  
 XX AAB10002;

XX AC AAB10002;

XX DT 01-NOV-2000 (first entry)

XX H. pylori 26 kDa protein-binding antibody heavy chain CDR2 peptide.

XX Acid-resistant microorganism; detection; faecal; intestine; infection;  
 KW monoclonal antibody; heavy chain; complementarity determining region;  
 KW CDR.

XX Unidentified.

XX WO200026671-A1.

XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-EP08212.

XX 29-OCT-1998; 98EP-0120517.

XX 06-NOV-1998; 98EP-0120687.

XX (CONN-) CONNEX GMBH.

XX Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;

XX Ringeis A;

XX WPI; 2000-365747/31.

XX N-PSDB; AAA40158.

XX Detecting infection by acid-fast microbes for diagnosis of Helicobacter

XX pylori, comprises reacting a faecal sample with two binding reagents for

XX Claim 22; Page 21; 84pp; German.

XX This invention describes a novel method for the detection of a mammalian  
 CC infection by an acid-resistant microorganism (A) by treating a faecal  
 CC sample with at least two different monoclonal antibodies (MAb) (or their  
 CC fragments or derivatives) or aptamers (collectively (I)) and detecting  
 CC formation of a complex (C) between (I) and the corresponding antigen of  
 CC (A). The first and second (I) bind to epitopes of different antigens  
 CC (Ag). These epitopes are present, after passage through the intestines,  
 CC in at least some mammals, and have either: (i) their native structure;  
 CC or (ii) a structure against which an antibody is produced by an animal  
 CC infected or immunized with (A), or its extract, lysate, derived protein  
 CC or fragment, or with a synthetic peptide. Practically all mammals display  
 CC at least one of the specified epitopes. The method is used to detect  
 CC infection by acid-fast bacteria, particularly of the genera Helicobacter,  
 CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,  
 CC M. tuberculosis, C. jejuni and C. pylori. (I) may also be used  
 CC therapeutically. The method is direct and non-invasive, and provides an  
 CC inexpensive and easily standardizable diagnosis, despite possible  
 CC degradation of antigens during passage through the intestines. This  
 CC sequence represents a fragment of a H. pylori 26 kDa protein-binding  
 CC antibody heavy chain complementarity determining region CDR2 which is  
 CC used to illustrate the method of the invention.

XX Sequence 19 AA;

Query Match 71.4%; Score 70; DB 21; Length 19;  
 Best Local Similarity 82.4%; Pred. No. 6.5e-05;  
 Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 IRLKSDNYATHYAESVK 18  
 |||:|||||

Db 2 IRLKSDNYATYANSVK 18  
 |||:|||||





PN WO9857994-A2.  
 XX  
 PD  
 XX  
 XX 23-DEC-1998.  
 XX  
 PF 16-JUN-1998; 98WO-US12402.  
 XX  
 XX 16-JUN-1997; 97US-0049871.  
 PR  
 XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.  
 PA  
 XX Fischer GW, Schuman RF, Stinson JL, Wong H;  
 PI WPI; 1999-095329/08.  
 XX N-PSDB; AAX05581.  
 DR  
 DR  
 XX New antibodies to lipoteichoic acid of gram positive bacteria - used  
 XX to develop products for the diagnosis, prevention and treatment of  
 PT infections caused by gram positive bacteria  
 PT  
 PT  
 XX Claim 21; Fig 12; 150pp; English.  
 PS  
 XX The invention relates to a monoclonal antibody (Mab) to lipoteichoic  
 CC acid of gram positive bacteria, where the Mab is a chimeric  
 CC immunoglobulin comprising at least part of a human immunoglobulin  
 CC constant region and at least part of a non-human immunoglobulin variable  
 CC region having specificity to lipoteichoic acid of gram positive bacteria.  
 CC The antibodies bind to whole bacteria and enhance phagocytosis and  
 CC killing of the bacteria and enhance protection from lethal infection. The  
 CC antibodies or peptides (encoded by a DNA of the variable region of  
 CC anti-lipoteichoic acid antibody or characterised by amino acids  
 CC corresponding to one or more of the complementarity determining regions  
 CC (CDRs) of the variable region of the antibody) can be used for treating  
 CC or preventing infections caused by gram positive bacteria. They can also  
 CC be used for the diagnosis of gram positive bacterial infections.  
 CC Sequences AAU94735-39 represent heavy chain variable regions of the  
 CC anti-lipoteichoic antibody 96-100.  
 XX  
 XX  
 SQ Sequence 19 AA;  
 Query Match 65.3%; Score 64; DB 20; Length 19;  
 Best Local Similarity 76.5%; Pred. NO. 0.0065;  
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Qy 2 IRLKSDNYATHYAESVK 18  
 |||||:|||||:  
 Db 2 IRLKSDNYATHYAESVK 18  
 |||||:|||||:  
 RESULT 9  
 AAU82539  
 ID AAU82539 standard; peptide; 17 AA.  
 XX  
 XX AAU82539;  
 AC  
 XX  
 XX 23-APR-2002 (first entry)  
 DT  
 XX Llama CDR2 region variable heavy chain fragment #32.  
 DE  
 XX Llama; phage display library; variable heavy domain fragment; VHH; VH;  
 KW sAb fragment; single domain anti-idiotypic antibody fragment;  
 KW phage display technology; immune system response; CDR1/H1; CDR2; CDR3;  
 KW complementarity determining region.  
 XX  
 OS Lama guanicoe glama.  
 XX  
 XX WO200190190-A2.  
 PN  
 XX 29-NOV-2001.  
 PD  
 XX 25-MAY-2001; 2001WO-CA00763.  
 PF  
 XX 26-MAY-2000; 2000US-207234P.  
 PR  
 XX

PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 XX Tanha J, Dubuc G, Narang S;  
 XX WPI; 2002-083093/11.  
 DR  
 XX New phage display library of variable heavy domain antigen-binding  
 PT fragments derived from llama antibodies, useful for in vitro selection  
 PT against any antigen of interest as a target -  
 XX  
 XX Claim 16; Page 32; 46pp; English.  
 PS  
 XX The present invention relates to a phage display library of variable  
 CC heavy domain (VHH or VH) fragments (sAb fragments) derived from llama  
 CC antibodies. The library is useful for in vitro selection against any  
 CC antigen of interest as a target. Single domain anti-idiotypic antibody  
 CC fragments are isolated from the library using phage display technology  
 CC and an antibody serving as an antigen. Such anti-idiotypic antibody  
 CC fragments have great potential in evoking the immune system response  
 CC to pathological antigens and in vaccine development. The large size of  
 CC the library considerably increases the probability of isolating from it  
 CC antigen-binding fragments having high affinity to almost any  
 CC predetermined target (antigen of interest). The library eliminates the  
 CC development of anti-idiotypic antibodies by immunisation and allows  
 CC isolation of anti-idiotypic antibodies. AAU82435-AAU82635 represent  
 CC the llama heavy chain domain fragments of the invention.  
 XX  
 XX Sequence 17 AA;  
 SQ  
 Query Match 53.1%; Score 52; DB 23; Length 17;  
 Best Local Similarity 58.8%; Pred. NO. 0.058;  
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 Qy 3 RLKSDNYATHYAESVK 19  
 |||||:|||||:  
 Db 1 RLKSDNYATHYAESVK 19  
 |||||:|||||:  
 RESULT 10  
 AAR74934  
 ID AAR74934 standard; peptide; 19 AA.  
 XX  
 XX AAR74934;  
 AC  
 XX  
 XX 19-JAN-1996 (first entry)  
 DT  
 XX H-CDR-2 of anti-idiotypic antibody against human anticancer antibody.  
 DE  
 XX Antibody; cancer; CDR; heavy chain; light chain; immunoglobulin;  
 KW complementarity determining region.  
 KW  
 XX Mus sp.  
 OS  
 XX JF07101999-A.  
 PN  
 XX 18-APR-1995.  
 PD  
 XX 06-OCT-1993; 93JP-0272950.  
 PF  
 XX 06-OCT-1993; 93JP-0272950.  
 PR  
 XX (HAGI/) HAGIWARA Y.  
 PA  
 XX WPI; 1995-182987/24.  
 DR  
 XX Novel anti-idiotypic antibody against an human anticancer monoclonal  
 PT antibody - and DNA sequences encoding the antibody, useful in  
 PT pharmacology, medicine and biochemical fields.  
 XX  
 XX Claim 1; Page 2; 28pp; Japanese.  
 PS  
 XX A new anti-idiotypic antibody against a human anticancer monoclonal  
 CC antibody is claimed. This antibody contains in its heavy chain 3  
 CC





XX  
PI Frewin MR, Gilliland LK, Tone M, Waldmann H, Walsh L;  
XX  
DR WPI; 1997-435093/40.  
XX  
PT Modified therapeutic antibody having reduced affinity for its  
XX antigen - can induce immunological tolerance, e.g. to Campath-1  
PT monoclonal antibodies used in antibody therapy of lymphoma and  
PT leukaemia  
XX  
PS Claim 12; Fig 1; 43pp; English.  
XX  
XX Peptides AM23390-95 represent "minimal" mutant sequences of the  
CC complementarity determining (CDR) region 2 of the humanised antibody  
CC Campath-1, Campath-1H. Campath-1H heavy chain sequence encodes human  
CC protein at all positions, except the 3 CDR regions. Campath-1 antibodies  
CC have been used to induce remissions in lymphoma and leukemia patients  
CC and for the treatment of rheumatoid arthritis and vasculitis. The target  
CC antigen is CD52, which is a GPI-anchored cell-surface glycoprotein of  
CC lymphocytes and monocytes. Loop 2, represented by residues 4-9 of the  
CC present sequence, is important for binding to CD52. Mutations were made  
CC within this region so as to abolish binding of the antibody to CD52,  
CC which reduces affinity for the antigen, and induces immunological  
CC tolerance. The present mutant contains a single charge difference from  
CC Lys to Asp. The positive charge of Lys is thought to interact with  
CC the negatively charged phosphate groups of the GPI anchor of CD52,  
CC and this single mutation may destroy antigen binding. Antibody  
CC therapy is often limited by an anti-antibody (antiglobulin) response  
CC preventing retreatment with the antibody. The minimal mutants or  
CC fragments may be used to induce tolerance to the therapeutic antibody in  
CC patients. They can also be used in the manufacture of medicines for  
CC induction of tolerance.  
XX  
SQ Sequence 19 AA;  
Query Match 45.9%; Score 45; DB 18; Length 19;  
Best Local Similarity 55.6%; Pred. No. 0.98;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
Oy 2 IRLKSDNYATHYAESVKG 19  
Db 2 IRLKADGYTTEYNPSVLG 19

Search completed: April 22, 2003, 13:48:26  
Job time : 74 secs

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:49:38 ; Search time 41 Seconds  
(without alignments)  
37.133 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98

Sequence: 1 EIRKSDNYATHYAESVKG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 53369

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pdb.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pdb.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pdb.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pdb.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pdb.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pdb.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pdb.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pdb.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pdb.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pdb.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pdb.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pdb.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	69.4	19	10	US-09-839-666-4
2	65	66.3	19	10	US-09-835-087-18
3	48	49.0	17	9	US-09-968-561A-98
4	48	49.0	17	10	US-09-192-854-60
5	46	46.9	17	9	US-09-968-561A-290
6	46	46.9	17	10	US-09-192-854-164
7	45	45.9	19	10	US-09-125-460A-4
8	45	45.9	19	10	US-09-730-857-85
9	43	43.9	17	9	US-09-968-561A-56
10	43	43.9	17	10	US-09-192-854-37
11	43	43.9	17	10	US-09-828-708-26
12	42	42.9	17	10	US-09-798-058-6
13	41	41.8	17	9	US-09-968-561A-8
14	41	41.8	17	9	US-09-968-561A-200
15	41	41.8	17	9	US-09-968-561A-206
16	41	41.8	17	10	US-09-192-854-4
17	41	41.8	17	10	US-09-192-854-115
18	41	41.8	17	10	US-09-125-460A-6
19	40	40.8	17	9	US-09-968-561A-14

20 40 40.8 17 9 US-09-968-561A-26  
21 40 40.8 17 9 US-09-968-561A-38  
22 40 40.8 17 9 US-09-968-561A-68  
23 40 40.8 17 9 US-09-968-561A-74  
24 40 40.8 17 9 US-09-968-561A-80  
25 40 40.8 17 9 US-09-968-561A-86  
26 40 40.8 17 9 US-09-968-561A-92  
27 40 40.8 17 9 US-09-968-561A-122  
28 40 40.8 17 9 US-09-968-561A-128  
29 40 40.8 17 9 US-09-968-561A-134  
30 40 40.8 17 9 US-09-968-561A-140  
31 40 40.8 17 9 US-09-968-561A-146  
32 40 40.8 17 9 US-09-968-561A-170  
33 40 40.8 17 9 US-09-968-561A-188  
34 40 40.8 17 9 US-09-968-561A-194  
35 40 40.8 17 9 US-09-968-561A-212  
36 40 40.8 17 9 US-09-968-561A-218  
37 40 40.8 17 9 US-09-968-561A-236  
38 40 40.8 17 9 US-09-968-561A-260  
39 40 40.8 17 9 US-09-968-561A-266  
40 40 40.8 17 9 US-09-968-561A-272  
41 40 40.8 17 9 US-09-968-561A-284  
42 40 40.8 17 9 US-09-968-561A-296  
43 40 40.8 17 9 US-09-968-561A-308  
44 40 40.8 17 9 US-09-968-561A-314  
45 40 40.8 17 9 US-10-161-145-13

#### ALIGNMENTS

RESULT 1  
US-09-839-666-4  
; Sequence 4, Application US/09839666  
; Patent No. US2002002513A1  
; GENERAL INFORMATION:  
; APPLICANT: SALLBERG, MATTI  
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY EXCHANGER  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DARBY & DARBY PC  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/839,666  
; FILING DATE: 19-Apr-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/737,085  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Reza  
; REGISTRATION NUMBER: 38,475  
; REFERENCE/DOCKET NUMBER: 3846/0C569  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-527-7659  
; TELEFAX: 212-753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

```

; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-839-666-4

Query Match          69.4%; Score 68; DB 10; Length 19;
Best Local Similarity 72.2%; Pred. No. 0.00017;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 IRLKSDNYATHVAESVK 19
   : ||| ||||| : |||||
Db 2 VRKSFNYATYADSVKG 19

RESULT 2
US-09-835-087-18
; Sequence 18, Application US/09835087
; Patent No. US20020042370A1
; GENERAL INFORMATION:
; APPLICANT: Wayne W. Hancock
; TITLE OF INVENTION: Method of Treating Graft Rejection Using
; TITLE OF INVENTION: Inhibitors of CCR2 Function
; FILE REFERENCE: 1855.2008-003
; CURRENT APPLICATION NUMBER: US/09/835,087
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/549,448
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)...(19)
; OTHER INFORMATION: CDR2 of murine mAb 1D9 heavy chain variable region
US-09-835-087-18

Query Match          66.3%; Score 65; DB 10; Length 19;
Best Local Similarity 70.6%; Pred. No. 0.00051;
Matches 12; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 IRLKSDNYATHVAESVK 18
   ||| : ||||| : |||||
Db 2 IRTKNNYATYADSVKG 18

RESULT 3
US-09-968-561A-98
; Sequence 98, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
;
Query Match          46.9%; Score 46; DB 9; Length 17;
Best Local Similarity 64.3%; Pred. No. 0.45;

```

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 6 SDNYATHYAESVKG 19  
Db 4 SSGYGTGYADSVKG 17

## RESULT 6

US-09-192-854-164  
; Sequence 164, Application US/09192854  
; Patent No. US20020068276A1  
; GENERAL INFORMATION:  
; APPLICANT: Tomlinson, Ian  
; TITLE OF INVENTION: Methods for Selecting Functional Peptides  
; CURRENT APPLICATION NUMBER: US/09/192,854  
; CURRENT FILING DATE: 1998-11-17  
; EARLIER FILING DATE: 1997-11-21  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 164  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-192-854-164

Query Match 46.9%; Score 46; DB 10; Length 17;  
Best Local Similarity 64.3%; Pred. No. 0.45;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 6 SDNYATHYAESVKG 19  
Db 4 SSGYGTGYADSVKG 17

## RESULT 7

US-09-125-460A-4  
; Sequence 4, Application US/09125460A  
; Patent No. US20020048578A1  
; GENERAL INFORMATION:  
; APPLICANT: WALDMANN, Herwan  
; APPLICANT: GILLILAND, Lisa K  
; APPLICANT: TONE, MASAHIDE  
; APPLICANT: FREWIN, MARK R  
; APPLICANT: WALSH, LOUISE  
; TITLE OF INVENTION: ANTIBODY VARIANTS  
; FILE REFERENCE: PP/NT/2794 US/09/125,460A  
; CURRENT APPLICATION NUMBER: US/09/125,460A  
; CURRENT FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: PCT/GB97/00472  
; PRIOR FILING DATE: 1997-02-20  
; PRIOR APPLICATION NUMBER: GB 9603507.6  
; PRIOR FILING DATE: 1996-02-20  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Peptide  
US-09-125-460A-4

Query Match 45.9%; Score 45; DB 10; Length 19;  
Best Local Similarity 55.6%; Pred. No. 0.73;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 IRLKSDNYATHYAESVKG 19  
Db 2 IRDKADGYTEYNPVSIG 19

## RESULT 8

US-09-730-857-85  
; Sequence 85, Application US/09730857  
; Patent No. US20020082396A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsushima, Kouji  
; Matsumoto, Yoshihiro  
; Yamada, Yoshiki  
; Sato, Koh  
; Tsuchiya, Masayuki  
; Tezuka, Tatsumi  
; TITLE OF INVENTION: Reshaped Human Antibody to  
; Interleukin-8  
; NUMBER OF SEQUENCES: 105  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1888  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/730,857  
; FILING DATE: 07-Dec-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/416,557  
; FILING DATE: 1999-10-12  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murashige, Kate H  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 35029-20001.10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-887-1500  
; TELEFAX: 202-822-0168  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 85:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 85:  
US-09-730-857-85

Query Match 45.9%; Score 45; DB 10; Length 19;  
Best Local Similarity 50.0%; Pred. No. 0.73;  
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IRLKSDNYATHYAESVKG 19  
Db 2 IRNKANGTYRETSASVKG 19

## RESULT 9

US-09-968-561A-56  
; Sequence 56, Application US/09968561A  
; Patent No. US20020164642A1  
; GENERAL INFORMATION:  
; APPLICANT: Tomlinson, Ian M  
; APPLICANT: Winter, Gregory  
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands  
; FILE REFERENCE: 8039/1073B  
; CURRENT APPLICATION NUMBER: US/09/968,561A  
; CURRENT FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: GB 9722131.1  
; PRIOR FILING DATE: 1997-10-20

```

; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-56

Query Match      43.9%; Score 43; DB 9; Length 17;
Best Local Similarity 47.1%; Pred. No. 1.3;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVKG 19
   | : : : : :
Db 1 RITPAGHRTYYADSVKG 17

RESULT 10
US-09-192-854-37
; Sequence 37, Application US/09192854
; Patent No. US20020068276A1
; GENERAL INFORMATION:
; APPLICANT: Winter, Greg
; TITLE OF INVENTION: Methods for Selecting Functional Peptides
; FILE REFERENCE: 3789/72916
; CURRENT APPLICATION NUMBER: US/09/192,854
; EARLIER APPLICATION NUMBER: 60/066,729
; PRIOR FILING DATE: 1998-11-17
; PRIOR FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-192-854-37

Query Match      43.9%; Score 43; DB 10; Length 17;
Best Local Similarity 47.1%; Pred. No. 1.3;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVKG 19
   | : : : : :
Db 1 RITPAGHRTYYADSVKG 17

RESULT 11
US-09-828-708-26
; Sequence 26, Application US/09828708
; Patent No. US20020146753A1
; GENERAL INFORMATION:
; APPLICANT: Ditzel, H.
; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their partic
; FILE REFERENCE: 1361.005U51
; CURRENT APPLICATION NUMBER: US/09/828,708
; PRIOR FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 17
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-09-828-708-26

Query Match      43.9%; Score 43; DB 10; Length 17;
Best Local Similarity 47.1%; Pred. No. 1.3;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 RLKSDNYATHYAESVKG 19
   | : : : : :
Db 1 RISGNSGTFYADSVKG 17

RESULT 12
US-09-798-058-6
; Sequence 6, Application US/09798058
; Patent No. US20020098523A1
; GENERAL INFORMATION:
; APPLICANT: Vaughan, Tristan John
; APPLICANT: Wilton, Alison Jane
; APPLICANT: Smith, Stephen
; APPLICANT: Main, Sarah Helen
; TITLE OF INVENTION: Human antibodies against eotaxin and their use
; FILE REFERENCE: 84632-000100
; CURRENT APPLICATION NUMBER: US/09/798,058
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 60/187,246
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-058-6

Query Match      42.9%; Score 42; DB 10; Length 17;
Best Local Similarity 61.5%; Pred. No. 1.9;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 7 DNYATHYAESVKG 19
   | : : : : :
Db 5 DGSIKHYADSVKG 17

RESULT 13
US-09-968-561A-8
; Sequence 8, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-8

```



```
Query Match      41.8%; Score 41; DB 9; Length 17;
Best Local Similarity 57.1%; Pred. No. 2.8;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 SDNYATHAESVKG 19
   |:|:|:|:|:|:|
Db 4 SEGWPITYADSVKG 17

RESULT 14
US-09-968-561A-200
; Sequence 200, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; APPLICANT: Winter, Gregory
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 200
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-200

Query Match      41.8%; Score 41; DB 9; Length 17;
Best Local Similarity 80.0%; Pred. No. 2.8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 ATHYAESVKG 19
   |:|:|:|:|
Db 8 ATRYADSVKG 17

RESULT 15
US-09-968-561A-206
; Sequence 206, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; APPLICANT: Winter, Gregory
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 206
; LENGTH: 17
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-968-561A-206

Query Match      41.8%; Score 41; DB 9; Length 17;
Best Local Similarity 80.0%; Pred. No. 2.8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 ATHYAESVKG 19
   |:|:|:|:|
Db 8 ATRYADSVKG 17

Search completed: April 22, 2003, 13:57:09
Job time : 42 secs
```



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 13:47:08 ; Search time 14 Seconds  
(without alignments)  
39.931 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98

Sequence: 1 EIRKSDNYATHYAESVKG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 127244

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	69.4	19	2	US-08-737-085A-4
2	68	69.4	19	3	US-09-246-258-4
3	68	69.4	19	4	US-09-532-106-4
4	68	69.4	19	4	US-09-839-666-4
5	66	67.3	19	1	US-07-977-696C-78
6	66	67.3	19	1	US-08-129-930B-78
7	66	67.3	19	4	US-08-976-288A-78
8	49	50.0	19	1	US-08-318-970B-6
9	47	48.0	19	1	US-08-318-970B-5
10	45	45.9	19	2	US-08-765-783A-85
11	45	45.9	19	4	US-09-416-557-85
12	44	44.9	17	1	US-08-264-093-22
13	39	39.8	14	2	US-08-452-724A-41
14	39	39.8	17	2	US-08-650-262-14
15	38	38.8	17	2	US-08-053-171-31
16	37	37.8	19	4	US-09-280-028-8
17	36	36.7	17	2	US-08-480-434-25
18	36	36.7	17	2	US-08-480-434-33
19	36	36.7	17	2	US-08-053-451B-25
20	36	36.7	17	2	US-08-053-451B-33
21	34	34.7	17	1	US-07-988-925-2
22	34	34.7	17	2	US-08-362-780-2
23	33	33.7	16	1	US-08-366-953A-30
24	33	33.7	17	1	US-08-244-626-14
25	33	33.7	19	2	US-08-480-434-34
26	33	33.7	19	2	US-08-053-451B-34
27	32	32.7	16	4	US-09-170-769A-11

28	32.7	17	1	US-08-285-936-54	Sequence 54, Appl
29	32.7	17	1	US-08-487-860-54	Sequence 54, Appl
30	31.6	16	1	US-08-208-886C-84	Sequence 84, Appl
31	31.6	16	1	US-08-704-744-86	Sequence 86, Appl
32	31.6	16	1	US-08-463-557-65	Sequence 65, Appl
33	31.6	16	2	US-08-290-793B-65	Sequence 65, Appl
34	30.6	14	6	5496552-4	Patent No. 5496552
35	30.6	17	4	US-09-406-532-16	Sequence 16, Appl
36	30.6	17	4	US-09-406-535-6	Sequence 6, Appl
37	29.6	8	6	5169835-28	Patent No. 5169835
38	29.6	16	1	US-08-208-886C-81	Sequence 81, Appl
39	29.6	16	1	US-08-208-886C-88	Sequence 88, Appl
40	29.6	16	1	US-08-704-744-83	Sequence 83, Appl
41	29.6	16	1	US-08-704-744-90	Sequence 90, Appl
42	29.6	16	1	US-08-366-953A-29	Sequence 29, Appl
43	29.6	16	1	US-08-366-953A-31	Sequence 31, Appl
44	29.6	16	1	US-08-469-557-62	Sequence 62, Appl
45	29.6	16	1	US-08-469-557-69	Sequence 69, Appl

## ALIGNMENTS

RESULT 1  
US-08-737-085A-4  
; Sequence 4, Application US/08737085A  
; Patent No. 5869232  
; GENERAL INFORMATION:  
; APPLICANT: SALLBERG, MATTI  
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DARBY & DARBY PC  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/737,085A  
; FILING DATE: 27-DEC-1996  
; CLASSIFICATION: 426  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Reza  
; REGISTRATION NUMBER: 38,475  
; REFERENCE/DOCKET NUMBER: 3846/OC569  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-527-7659  
; TELEFAX: 212-753-6237  
; TELEX: 236687

\* INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-737-085A-4

Query Match 69.4%; Score 68; DB 2; Length 19;  
Best Local Similarity 72.2%; Pred. NO. 8.6e-05;  
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 IRLKSDNYATHYAESVKG 19

```

; :| || |||||:|||
Db      2 VRKSFNYATYADSVKG 19

RESULT 2
US-09-246-258-4
; Sequence 4, Application US/09246258
; Patent No. 6040137
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/246,258
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/737,085
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/OC569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-532-106-4
Query Match          69.4%; Score 68; DB 4; Length 19;
Best Local Similarity 72.2%; Pred. No. 8.6e-05;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      2 IRLKSDNYATHYAESVKG 19
       :| || |||||:|||
Db      2 VRKSFNYATYADSVKG 19

RESULT 4
US-09-839-666-4
; Sequence 4, Application US/09839666
; Patent No. 6469143
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/839,666
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/737,085
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/OC569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-246-258-4
Query Match          69.4%; Score 68; DB 3; Length 19;
Best Local Similarity 72.2%; Pred. No. 8.6e-05;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      2 IRLKSDNYATHYAESVKG 19
       :| || |||||:|||
Db      2 VRKSFNYATYADSVKG 19

RESULT 3
US-09-532-106-4
; Sequence 4, Application US/09532106
; Patent No. 6245895
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA

```

TELEFAX: 212-753-6237  
 TELEX: 236687  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 19 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
 US-09-839-666-4

Query Match 69.4%; Score 68; DB 4; Length 19;  
 Best Local Similarity 72.2%; Pred. No. 8.6e-05;  
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IRLKSDNYATHYAEVSKG 19

Db 2 VRSKSFNYATYADSVKG 19

## RESULT 5

US-07-977-696C-78  
 Sequence 78, Application US/07977696C  
 Patent No. 5792852  
 GENERAL INFORMATION:  
 APPLICANT: do Couto, Fernando J.R.  
 APPLICANT: Ceriani Dr., Roberto L.  
 APPLICANT: Peterson Dr., Jerry A.  
 APPLICANT: Padlan Dr., Eduardo A.  
 TITLE OF INVENTION: Analogue Peptides with Specificity  
 TITLE OF INVENTION: for Carcinomas and Kit and Diagnostic Vaccination  
 TITLE OF INVENTION: Therapeutic Methods.  
 NUMBER OF SEQUENCES: 81  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI  
 STREET: 444 South Flower Street, Suite 2000  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: USA  
 ZIP: 90071

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS 5.0  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/977,696C  
 FILING DATE: 11-16-92  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Amzel Ph.D., Viviana  
 REGISTRATION NUMBER: 30,930  
 REFERENCE/DOCKET NUMBER: P66 38227  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (510) 748-6868  
 TELEFAX: (510) 748-6688  
 TELEX: n.a.

INFORMATION FOR SEQ ID NO: 78:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 19 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-07-977-696C-78

Query Match 67.3%; Score 66; DB 1; Length 19;  
 Best Local Similarity 68.4%; Pred. No. 0.00018;  
 Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EIRLKSNDYATHYAEVSKG 19

Db 1 EIRKNNHATYDESVKG 19

## RESULT 6

US-08-129-930B-78  
 Sequence 78, Application US/08129930B  
 Patent No. 5804187  
 GENERAL INFORMATION:  
 APPLICANT: do Couto Dr., Fernando J.R.  
 APPLICANT: Ceriani Dr., Roberto L.  
 APPLICANT: Peterson Dr., Jerry A.  
 APPLICANT: Padlan Dr., Eduardo A.  
 TITLE OF INVENTION: Analogue Peptides with Broad  
 TITLE OF INVENTION: Carcinoma Specificity, and Kit and  
 TITLE OF INVENTION: Diagnostic Vaccination and  
 TITLE OF INVENTION: Therapeutic Methods  
 NUMBER OF SEQUENCES: 96  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: V. AMZEL & ASSOC.  
 STREET: 2055 No. 5804187th Broadway, Suite 201  
 CITY: Walnut Creek  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94596

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS 5.0  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/129,930B  
 FILING DATE: September 30, 1993  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Amzel Ph.D., Viviana  
 REGISTRATION NUMBER: 30,930  
 REFERENCE/DOCKET NUMBER: CRFCC-008A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (510) 521-1333  
 TELEFAX: (510) 521-3541  
 TELEX: n.a.  
 INFORMATION FOR SEQ ID NO: 78:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 19 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-129-930B-78

Query Match 67.3%; Score 66; DB 1; Length 19;  
 Best Local Similarity 68.4%; Pred. No. 0.00018;  
 Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EIRLKSNDYATHYAEVSKG 19

Db 1 EIRKNNHATYDESVKG 19

## RESULT 7

US-08-976-288A-78  
 Sequence 78, Application US/08976288A  
 Patent No. 6315997  
 GENERAL INFORMATION:  
 APPLICANT: do Couto Dr., Fernando J.R.  
 APPLICANT: Ceriani Dr., Roberto L.  
 APPLICANT: Peterson Dr., Jerry A.  
 APPLICANT: Padlan Dr., Eduardo A.  
 TITLE OF INVENTION: Analogue Peptides with Broad  
 TITLE OF INVENTION: Carcinoma Specificity, and Kit and  
 TITLE OF INVENTION: Diagnostic Vaccination and  
 TITLE OF INVENTION: Therapeutic Methods  
 NUMBER OF SEQUENCES: 96  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pretty, Schroeder & Poplawski

```

; STREET: 444 South Flower St., 19th Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
; SOFTWARE: Patent in Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,288A
; FILING DATE: No. 6315997ember 21, 1997
; CLASSIFICATION: 424
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/129,930
; FILING DATE: September 30, 1993
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,696
; FILING DATE: No. 6315997ember 16, 1992
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel Ph.D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: P6639938
;
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
;
; TELE: n.a.
;
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
; US-08-976-288A-78
;
; Query Match 67.3%; Score 66; DB 4; Length 19;
; Best Local Similarity 68.4%; Pred. No. 0.00018;
; Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
;
; Qy 1 BIRKSDNYATHYAESVKG 19
; Db 1 BIRKANNHATYDESXVG 19
;
; RESULT 8
; US-08-318-970B-6
; Sequence 6, Application US/08318970B
; Patent No. 5589573
;
; GENERAL INFORMATION:
; APPLICANT: Hideaki HAGIWARA, et al.
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF ANTI-IDIOTYPIC
; TITLE OF INVENTION: ANTIBODIES AGAINST ANTI-CANCER HUMAN MONOCLONAL ANTIBODY
; TITLE OF INVENTION: AND DNA BASE SEQUENCES ENCODING THOSE SEQUENCES
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Office of Sherman and Shalloway
; STREET: 413 N. Washington Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: Dell System 210; Intel 80 285 Microprocessor
; OPERATING SYSTEM: MS DOS 3.3
; SOFTWARE: Word Perfect, Version 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,970B
; FILING DATE: October 6, 1994
; CLASSIFICATION: 530
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard A. Steinberg
; REGISTRATION NUMBER: 26,588
; REFERENCE/DOCKET NUMBER: S-2371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 549-2282
; TELEFAX: (703) 836-0106
;
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: H-CDR2-2
; OTHER INFORMATION: hypervariable region
;
; US-08-318-970B-5
;
; Query Match 48.0%; Score 47; DB 1; Length 19;
; Best Local Similarity 55.6%; Pred. No. 0.21;
; Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard A. Steinberg

```

```

;
; REGISTRATION NUMBER: 26,588
; REFERENCE/DOCKET NUMBER: S-2371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 549-2282
; TELEFAX: (703) 836-0106
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: H-CDR2-3
; OTHER INFORMATION: hypervariable region
;
; US-08-318-970B-6
;
; Query Match 50.0%; Score 49; DB 1; Length 19;
; Best Local Similarity 55.6%; Pred. No. 0.1;
; Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
;
; Qy 2 IRLKSDNYATHYAESVKG 19
; Db 2 IRLKANNHATYDESXVG 19
;
; RESULT 9
; US-08-318-970B-5
; Sequence 5, Application US/08318970B
; Patent No. 5589573
;
; GENERAL INFORMATION:
; APPLICANT: Hideaki HAGIWARA, et al.
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF ANTI-IDIOTYPIC
; TITLE OF INVENTION: ANTIBODIES AGAINST ANTI-CANCER HUMAN MONOCLONAL ANTIBODY
; TITLE OF INVENTION: AND DNA BASE SEQUENCES ENCODING THOSE SEQUENCES
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Office of Sherman and Shalloway
; STREET: 413 N. Washington Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: Dell System 210; Intel 80 285 Microprocessor
; OPERATING SYSTEM: MS DOS 3.3
; SOFTWARE: Word Perfect, Version 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,970B
; FILING DATE: October 6, 1994
; CLASSIFICATION: 530
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard A. Steinberg
; REGISTRATION NUMBER: 26,588
; REFERENCE/DOCKET NUMBER: S-2371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 549-2282
; TELEFAX: (703) 836-0106
;
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: H-CDR2-2
; OTHER INFORMATION: hypervariable region
;
; US-08-318-970B-5
;
; Query Match 48.0%; Score 47; DB 1; Length 19;
; Best Local Similarity 55.6%; Pred. No. 0.21;
; Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard A. Steinberg

```



```

; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA: No. 5639863 applicable
; ATTORNEY/AGENT INFORMATION:
; NAME: Lake, James R. 31081
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: NOVOP/106A/7551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 868-1482
; TELEFAX: (416) 362-0823
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: not applicable
; TOPOLOGY: linear
US-08-264-093-22

Query Match 44.9%; Score 44; DB 1; Length 17;
Best Local Similarity 50.0%; Pred. No. 0.57;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 LKSDNYATHYAESVKG 19
Db 2 ISSNGSGTYADSVKG 17

RESULT 13
US-08-452-724A-41
; Sequence 41, Application US/08452724A
; Patent No. 5830850
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: Walk-Through Mutagenesis
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,724A
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,600
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION NUMBER: PCT/US91/02362
; APPLICATION NUMBER: US 07/505,314
; FILING DATE: 05-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook Esq., David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: RC90-01AZ
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown

```

```

US-08-452-724A-41

Query Match 39.8%; Score 39; DB 2; Length 14;
Best Local Similarity 53.8%; Pred. No. 3;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 DNYATHYAESVKG 19
Db 2 NKYTEYSASVKG 14

RESULT 14
US-08-650-262-14
; Sequence 14, Application US/08650262
; Patent No. 5866124
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 5866124man
; APPLICANT: Pluschke, Gerd
; APPLICANT: Murray, Brendan
; TITLE OF INVENTION: Genetically engineered antibodies
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5866124artis Corporation
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: NJ
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/650,262
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 92810189.0
; FILING DATE: 17-MAR-1992
; * ATTORNEY/AGENT INFORMATION:
; NAME: Henry P. No. 5866124ak
; REGISTRATION NUMBER: 33200
; REFERENCE/DOCKET NUMBER: 4-18995/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-277-5110
; TELEFAX: 908-277-4306
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-08-650-262-14

Query Match 39.8%; Score 39; DB 2; Length 17;
Best Local Similarity 43.8%; Pred. No. 3.7;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 LKSDNYATHYAESVKG 19
Db 2 ISSDSSNIYADTVKG 17

RESULT 15
US-08-053-171-31
; Sequence 31, Application US/08053171
; Patent No. 5562903
; GENERAL INFORMATION:
; APPLICANT: Co, Loibner
; TITLE OF INVENTION: Antibody Derivatives
; NUMBER OF SEQUENCES: 32

```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/053,171
FILING DATE: 22-APR-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-54-1
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..17
OTHER INFORMATION: /note= "Second
OTHER INFORMATION: complementarity-determining region (CDR2) of
OTHER INFORMATION: BR55-2 antibody heavy chain"
US-08-053-171-31

Query Match 38.8%; Score 38; DB 1; Length 17;
Best Local Similarity 60.0%; Pred. No. 5.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 10 ATHYAESVKG 19
Db 8 SSHYVDSVKG 17

Search completed: April 22, 2003, 13:50:14
Job time : 15 secs

```



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:49 ; Search time 16.625 Seconds  
(without alignments)  
33.626 Million cell updates/sec

Title: US-09-674-716B-11

Perfect score: 98

Sequence: 1 EIRKSDNYATHYAESVKG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgm2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgm2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgm2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgm2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgm2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	285	4	US-09-318-661-4
2	95	96.9	298	4	US-09-318-661-2
3	93	94.9	115	1	US-08-468-661-1
4	93	94.9	115	1	US-08-466-272A-1
5	93	94.9	115	1	US-08-478-857-1
6	93	94.9	115	2	US-08-471-771-1
7	93	94.9	115	3	US-09-130-783-1
8	93	94.9	120	3	US-08-767-128-28
9	93	94.9	122	3	US-08-483-749A-2
10	85	86.7	114	3	US-08-483-749A-10
11	80.5	82.1	119	3	US-08-767-128-26
12	74	75.5	119	1	US-08-192-102-5
13	74	75.5	119	1	US-08-324-799-5
14	74	75.5	119	2	US-08-192-861A-5
15	74	75.5	119	4	US-09-133-119-5
16	74	75.5	119	4	US-08-192-093A-5
17	72	73.5	119	1	US-08-442-542-45
18	72	73.5	119	3	US-08-765-469-45
19	71	72.4	227	1	US-08-681-432-2
20	68	69.4	19	2	US-08-737-085A-4
21	68	69.4	19	3	US-09-246-258-4
22	68	69.4	19	4	US-09-532-106-4
23	68	69.4	19	4	US-09-839-666-4
24	68	69.4	27	3	US-08-737-085A-21
25	68	69.4	27	3	US-09-246-258-21
26	68	69.4	27	4	US-09-532-106-21
27	68	69.4	27	4	US-09-839-666-21

28 67 68.4 110 3 US-08-767-128-24 Sequence 24, Appl  
29 67 68.4 115 3 US-08-767-128-36 Sequence 36, Appl  
30 66 67.3 19 1 US-07-977-696C-78 Sequence 78, Appl  
31 66 67.3 19 1 US-08-129-930B-78 Sequence 78, Appl  
32 66 67.3 19 4 US-08-976-288A-78 Sequence 13, Appl  
33 66 67.3 134 1 US-07-977-696C-13 Sequence 13, Appl  
34 66 67.3 134 1 US-08-129-930B-13 Sequence 13, Appl  
35 66 67.3 134 4 US-08-976-288A-13 Sequence 13, Appl  
36 64 65.3 119 1 US-08-442-542-14 Sequence 14, Appl  
37 64 65.3 119 3 US-08-765-469-14 Sequence 14, Appl  
38 60 61.2 119 1 US-08-442-542-2 Sequence 2, Appl  
39 60 61.2 119 3 US-08-765-469-2 Sequence 2, Appl  
40 60 61.2 599 1 US-08-442-542-18 Sequence 18, Appl  
41 60 61.2 599 3 US-08-765-469-18 Sequence 18, Appl  
42 58 59.2 125 2 US-08-428-197-9 Sequence 9, Appl  
43 58 59.2 125 5 PCT-US93-10555-9 Sequence 9, Appl  
44 57 58.2 35 4 US-08-525-539A-33 Sequence 33, Appl  
45 57 58.2 109 1 US-07-942-245-21 Sequence 21, Appl

## ALIGNMENTS

RESULT 1  
US-09-318-661-4  
; Sequence 4, Application US/09318661  
; Patent No. 6268488  
; GENERAL INFORMATION:  
; APPLICANT: Barbas III, Carlos F.  
; APPLICANT: Shabat, Doron  
; APPLICANT: Rader, Christoph  
; APPLICANT: List, Benjamin  
; APPLICANT: Lerner, Richard A.  
; TITLE OF INVENTION: PRODRUG ACTIVATION USING CATALYTIC ANTIBODIES  
; FILE REFERENCE: PLF00115  
; CURRENT APPLICATION NUMBER: US/09/318,661  
; CURRENT FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4  
; TYPE: PRT  
; LENGTH: 285  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Nucleotide  
; OTHER INFORMATION: residue sequence of catalytic fragment  
US-09-318-661-4

Query Match 100.0%; Score 98; DB 4; Length 285;  
Best Local Similarity 100.0%; Pred. No. 2.4e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRKSDNYATHYAESVKG 19  
|||||  
DB 203 EIRKSDNYATHYAESVKG 221

RESULT 2  
US-09-318-661-2  
; Sequence 2, Application US/09318661  
; Patent No. 6268488  
; GENERAL INFORMATION:  
; APPLICANT: Barbas III, Carlos F.  
; APPLICANT: Shabat, Doron  
; APPLICANT: Rader, Christoph  
; APPLICANT: List, Benjamin  
; APPLICANT: Lerner, Richard A.  
; TITLE OF INVENTION: PRODRUG ACTIVATION USING CATALYTIC ANTIBODIES  
; FILE REFERENCE: PLF00115  
; CURRENT APPLICATION NUMBER: US/09/318,661  
; CURRENT FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin Ver. 2.1

```
; SEQ ID NO 2
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-318-661-2

Query Match          96.9%; Score 95; DB 4; Length 298;
Best Local Similarity 94.7%; Pred. No. 7.8e-08;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EIRKSDNYATHYAESVKG 19
|||||:|||||:|||||:
Db 205 EIRKSDNYATHYAESVKG 223

RESULT 3
US-08-468-661-1
; Sequence 1, Application US/08468661
; Patent No. 5639621
; GENERAL INFORMATION:
; APPLICANT: Boselet, Klaus
; APPLICANT: Pfeleiderer, Peter
; APPLICANT: Seeman, Gerhard
; TITLE OF INVENTION: Monoclonal Antibodies Against
; TITLE OF INVENTION: Tumor-Associated Antigens, Processes for the
; TITLE OF INVENTION: Preparation Thereof and the Use Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,661
; FILING DATE: 06-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/957,827
; FILING DATE: 08-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Forman, David S.
; REGISTRATION NUMBER: 35,694
; REFERENCE/DOCKET NUMBER: 05552-1227-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 115 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-468-661-1

Query Match          94.9%; Score 93; DB 1; Length 115;
Best Local Similarity 94.7%; Pred. No. 5.7e-08;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EIRKSDNYATHYAESVKG 19
|||||:|||||:|||||:
Db 46 EIRKSDNYATHYAESVKG 64

RESULT 5
US-08-478-857-1
; Sequence 1, Application US/08478857
; Patent No. 5695758
; GENERAL INFORMATION:
; APPLICANT: Boselet, Klaus
; APPLICANT: Pfeleiderer, Peter
; APPLICANT: Seeman, Gerhard
; TITLE OF INVENTION: Monoclonal Antibodies Against
; TITLE OF INVENTION: Tumor-Associated Antigens, Processes for the
; TITLE OF INVENTION: Preparation Thereof and the Use Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,661
; FILING DATE: 06-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/957,827
; FILING DATE: 08-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Forman, David S.
; REGISTRATION NUMBER: 35,694
; REFERENCE/DOCKET NUMBER: 05552-1227-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 115 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-468-661-1

Query Match          94.9%; Score 93; DB 1; Length 115;
Best Local Similarity 94.7%; Pred. No. 5.7e-08;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EIRKSDNYATHYAESVKG 19
|||||:|||||:|||||:
Db 46 EIRKSDNYATHYAESVKG 64

RESULT 4
US-08-468-661-1
; Sequence 1, Application US/08466272A
; Patent No. 5674994
; GENERAL INFORMATION:
; APPLICANT: Boselet, Klaus
; APPLICANT: Pfeleiderer, Peter
; APPLICANT: Seeman, Gerhard
; TITLE OF INVENTION: Monoclonal Antibodies Against
; TITLE OF INVENTION: Tumor-Associated Antigens, Processes for the
; TITLE OF INVENTION: Preparation Thereof and the Use Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,272A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/957,827
; FILING DATE: 08-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 02481-1227-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 115 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-466-272A-1

Query Match          94.9%; Score 93; DB 1; Length 115;
Best Local Similarity 94.7%; Pred. No. 5.7e-08;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EIRKSDNYATHYAESVKG 19
|||||:|||||:|||||:
Db 46 EIRKSDNYATHYAESVKG 64

RESULT 5
US-08-478-857-1
; Sequence 1, Application US/08478857
; Patent No. 5695758
; GENERAL INFORMATION:
; APPLICANT: Boselet, Klaus
; APPLICANT: Pfeleiderer, Peter
; APPLICANT: Seeman, Gerhard
; TITLE OF INVENTION: Monoclonal Antibodies Against
; TITLE OF INVENTION: Tumor-Associated Antigens, Processes for the
; TITLE OF INVENTION: Preparation Thereof and the Use Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
```

COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,857  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/957,827  
FILING DATE: 08-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bugalsky, Lawrence B.  
REGISTRATION NUMBER: 35,086  
REFERENCE/DOCKET NUMBER: 02481-1227-00000  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-478-857-1

Query Match 94.9%; Score 93; DB 1; Length 115;  
Best Local Similarity 94.7%; Pred. No. 5.7e-08;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRKSDNYATHYAESVKG 19  
|||||:|||||

Db 46 EIRKSNYATHYAESVKG 64

## RESULT 6

US-08-471-771-1  
Sequence 1, Application US/08/471771  
Patent No. 5837824  
GENERAL INFORMATION:  
APPLICANT: Bosslet, Klaus  
APPLICANT: Pfeleiderer, Peter  
APPLICANT: Seeman, Gerhard  
TITLE OF INVENTION: Monoclonal Antibodies Against  
Tumor-Associated Antigens, Processes for the  
Preparation Thereof and the Use Thereof  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,771  
FILING DATE: 06-JUNE-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/957,827  
FILING DATE: 08-OCT-1992  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Forman, David S.

REGISTRATION NUMBER: 33,694  
REFERENCE/DOCKET NUMBER: 05552-1227-01000  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-471-771-1

Query Match 94.9%; Score 93; DB 2; Length 115;  
Best Local Similarity 94.7%; Pred. No. 5.7e-08;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRKSDNYATHYAESVKG 19

|||||:|||||

Db 46 EIRKSNYATHYAESVKG 64

## RESULT 7

US-09-130-783-1  
Sequence 1, Application US/09130783  
Patent No. 6030797  
GENERAL INFORMATION:  
APPLICANT: Bosslet, Klaus  
APPLICANT: Pfeleiderer, Peter  
APPLICANT: Seeman, Gerhard  
TITLE OF INVENTION: Monoclonal Antibodies Against  
Tumor-Associated Antigens, Processes for the  
Preparation Thereof and the Use Thereof  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/130,783  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/471,771  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Forman, David S.  
REGISTRATION NUMBER: 33,694  
REFERENCE/DOCKET NUMBER: 05552-1227-01000  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-130-783-1

Query Match 94.9%; Score 93; DB 3; Length 115;  
Best Local Similarity 94.7%; Pred. No. 5.7e-08;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRKSDNYATHYAESVKG 19  
|||||:|||||  
Db 46 EIRKSDNYATHYAESVKG 64

RESULT 8  
US-08-767-128-28  
; Sequence 28, Application US/08767128  
; Patent No. 6111079  
; GENERAL INFORMATION:  
; APPLICANT: WYLIE, DWANE E.  
; APPLICANT: LOPEZ, OSVALDO  
; APPLICANT: MURRAY, PETER JOSEPH  
; APPLICANT: GOEBEL, PETER  
; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
; TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
; STREET: 3100 No. 6111079west Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/767,128  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09258  
; FILING DATE: 05-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/541,373  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/462,798  
; FILING DATE: 05-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carter, Charles G.  
; REGISTRATION NUMBER: 35,093  
; REFERENCE/DOCKET NUMBER: 8648.49USF1  
; TELEPHONE: 612/371-5278  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-08-767-128-28

Query Match 94.9%; Score 93; DB 3; Length 120;  
Best Local Similarity 94.7%; Pred. No. 6e-08;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRKSDNYATHYAESVKG 19

Db 50 EIRKSDNYATHYAESVKG 68  
|||||:|||||  
RESULT 9  
US-08-483-749A-2  
; Sequence 2, Application US/08483749A  
; Patent No. 6054561  
; GENERAL INFORMATION:  
; APPLICANT: RING, DAVID B.  
; TITLE OF INVENTION: ANTIGEN-BINDING SITES OF ANTIBODY  
; TITLE OF INVENTION: MOLECULES SPECIFIC FOR CANCER ANTIGENS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CHIRON CORPORATION  
; STREET: INTELLECTUAL PROPERTY - R440, PO BOX 8097  
; CITY: EMERYVILLE  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,749A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAVERIDE, PAUL B.  
; REGISTRATION NUMBER: 36,914  
; REFERENCE/DOCKET NUMBER: 0508.008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2585  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 122 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-483-749A-2

Query Match 94.9%; Score 93; DB 3; Length 122;  
Best Local Similarity 94.7%; Pred. No. 6.1e-08;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRKSDNYATHYAESVKG 19

Db 50 EIRKSDNYATHYAESVKG 68  
|||||:|||||

RESULT 10  
US-08-483-749A-10  
; Sequence 10, Application US/08483749A  
; Patent No. 6054561  
; GENERAL INFORMATION:  
; APPLICANT: RING, DAVID B.  
; TITLE OF INVENTION: ANTIGEN-BINDING SITES OF ANTIBODY  
; TITLE OF INVENTION: MOLECULES SPECIFIC FOR CANCER ANTIGENS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CHIRON CORPORATION  
; STREET: INTELLECTUAL PROPERTY - R440, PO BOX 8097  
; CITY: EMERYVILLE  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/483,749A  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: SAVERIDE, PAUL B.  
 REGISTRATION NUMBER: 36,914  
 REFERENCE/DOCKET NUMBER: 0508.008  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (510) 601-2585  
 TELEFAX: (510) 655-3542  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 114 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-483-749A-10

Query Match 86.7%; Score 85; DB 3; Length 114;  
 Best Local Similarity 84.2%; Pred. No. 1.1e-06;  
 Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRLKSDNYATHYAESVKG 19  
 |||:|||||  
 Db 50 EIRLKSDNYATHYAESVKG 68

RESULT 11  
 US-08-767-128-26  
 Sequence 26, Application US/08/67128  
 Patent No. 611079  
 GENERAL INFORMATION:  
 APPLICANT: WYLIE, DWANE E.  
 APPLICANT: LOPEZ, OSVALDO  
 APPLICANT: MURRAY, PETER JOSEPH  
 APPLICANT: GORBEL, PETER  
 TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
 TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
 NUMBER OF SEQUENCES: 46  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
 STREET: 3100 No. 611079west Center, 90 South Seventh St  
 CITY: Minneapolis  
 STATE: MN  
 COUNTRY: USA  
 ZIP: 55402  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/767,128  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE: 04-DEC-1996  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US96/09258  
 FILING DATE: 05-JUN-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/541,373  
 FILING DATE: 10-OCT-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/462,798  
 FILING DATE: 05-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Carter, Charles G.  
 REGISTRATION NUMBER: 35,093

REFERENCE/DOCKET NUMBER: 8648.49USFT1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 612/371-5278  
 TELEFAX: 612/332-9081  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 26:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 119 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: internal  
 ORIGINAL SOURCE:  
 US-08-767-128-26

Query Match 82.1%; Score 80.5; DB 3; Length 119;  
 Best Local Similarity 89.5%; Pred. No. 6.2e-06;  
 Matches 17; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 EIRLKSDNYATHYAESVKG 19  
 |||:|||||  
 Db 50 EIRLKSDNYATHYAESVKG 67

RESULT 12  
 US-08-192-102-5  
 Sequence 5, Application US/08192102  
 Patent No. 5656272  
 GENERAL INFORMATION:  
 APPLICANT: Le, Junning  
 APPLICANT: Vilecek, Jan  
 APPLICANT: Daddona, Peter E.  
 APPLICANT: Ghayeb, John  
 APPLICANT: Knight, David M.  
 APPLICANT: Siegel, Scott A.  
 TITLE OF INVENTION: ANTI-TNF ANTIBODIES AND ASSAYS EMPLOYING  
 TITLE OF INVENTION: ANTI-TNF ANTIBODIES  
 NUMBER OF SEQUENCES: 19  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/192,102  
 FILING DATE: 04-FEB-1994  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/192,093  
 FILING DATE: 04-FEB-1994  
 APPLICATION NUMBER: US 08/013,413  
 FILING DATE: 02-FEB-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/010,406  
 FILING DATE: 29-JAN-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/943,852  
 FILING DATE: 11-SEP-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/853,606  
 FILING DATE: 18-MAR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/670,827

```

; FILING DATE: 18-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: NYU93-01M3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-192-102-5

Query Match 75.5%; Score 74; DB 1; Length 119;
Best Local Similarity 84.2%; Pred. No. 7e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EIRKSDNYATHYAESVKG 19
Db 50 EIRKSDNYATHYAESVKG 68

RESULT 13
US-08-324-799-5
; Sequence 5, Application US/08324799
; Patent No. 5698195
; GENERAL INFORMATION:
; APPLICANT: Le, Junning
; APPLICANT: Vilcek, Jan
; APPLICANT: Daddona, Peter E.
; APPLICANT: Ghayeb, John
; APPLICANT: Knight, David M.
; APPLICANT: Siegel, Scott A.
; TITLE OF INVENTION: ANTI-TNF ANTIBODIES AND PEPTIDES
; TITLE OF INVENTION: OF HUMAN TUMOR NECROSIS FACTOR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,799
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,093
; FILING DATE: 04-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,102
; FILING DATE: 04-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,861
; FILING DATE: 04-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/013,413
; FILING DATE: 02-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,406
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,852
; FILING DATE: 11-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/670,827
; FILING DATE: 18-MAR-1991

; APPLICATION NUMBER: US 07/853,606
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/670,827
; FILING DATE: 18-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: NYU93-01M4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-324-799-5

Query Match 75.5%; Score 74; DB 1; Length 119;
Best Local Similarity 84.2%; Pred. No. 7e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EIRKSDNYATHYAESVKG 19
Db 50 EIRKSDNYATHYAESVKG 68

RESULT 14
US-08-192-861A-5
; Sequence 5, Application US/08192861A
; Patent No. 5919452
; GENERAL INFORMATION:
; APPLICANT: Le, Junning
; APPLICANT: Vilcek, Jan
; APPLICANT: Daddona, Peter E.
; APPLICANT: Ghayeb, John
; APPLICANT: Knight, David M.
; APPLICANT: Siegel, Scott A.
; TITLE OF INVENTION: METHODS OF TREATING TNF-MEDIATED DISEASE USING
; TITLE OF INVENTION: CHIMERIC ANTI-TNF ANTIBODIES (As Amended)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/192,861A
; FILING DATE: 04-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/013,413
; FILING DATE: 02-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,406
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,852
; FILING DATE: 11-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/853,606
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/670,827
; FILING DATE: 18-MAR-1991
```







GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:49 Search time 7.875 seconds  
(without alignments)  
33.626 Million cell updates/sec

Title: US-09-674-716B-7  
Perfect score: 48  
Sequence: 1 QQLVEYPFT 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	91.7	9	2	US-08-672-345C-27
2	44	91.7	9	4	US-09-214-095D-27
3	44	91.7	113	2	US-08-672-345C-8
4	44	91.7	113	2	US-08-672-345C-98
5	44	91.7	113	4	US-09-214-095D-8
6	44	91.7	113	4	US-09-214-095D-108
7	36	75.0	9	2	US-08-672-345C-21
8	36	75.0	9	4	US-09-214-095D-21
9	36	75.0	113	2	US-08-672-345C-5
10	36	75.0	113	2	US-08-672-345C-7
11	36	75.0	113	2	US-08-672-345C-95
12	36	75.0	113	4	US-09-214-095D-5
13	36	75.0	113	4	US-09-214-095D-7
14	36	75.0	113	4	US-09-214-095D-100
15	36	75.0	113	4	US-09-214-095D-112
16	36	75.0	113	4	US-09-214-095D-119
17	36	75.0	280	4	US-09-214-095D-112
18	35	72.9	112	3	US-08-483-749A-4
19	34	70.8	9	4	US-09-406-532-20
20	34	70.8	113	4	US-09-406-532-14
21	34	70.8	585	4	US-09-134-001C-4914
22	33	68.8	9	1	US-08-438-123-3
23	33	68.8	11	1	US-08-438-123-11
24	33	68.8	107	4	US-09-240-274-40
25	33	68.8	108	1	US-08-468-661-3
26	33	68.8	108	1	US-08-466-272A-3
27	33	68.8	108	1	US-08-478-857-3

28	33	68.8	108	2	US-08-471-771-3	Sequence 3, Appli
29	33	68.8	108	3	US-09-130-783-3	Sequence 3, Appli
30	33	68.8	113	3	US-08-483-749A-16	Sequence 16, Appli
31	33	68.8	113	5	PCT-US93-11611-5	Sequence 5, Appli
32	33	68.8	133	1	US-08-253-877C-10	Sequence 10, Appli
33	33	68.8	133	1	US-08-253-877C-28	Sequence 28, Appli
34	33	68.8	133	2	US-08-452-164A-10	Sequence 10, Appli
35	33	68.8	133	2	US-08-452-164A-28	Sequence 28, Appli
36	33	68.8	133	3	US-08-603-024-4	Sequence 4, Appli
37	33	68.8	133	3	US-08-603-024-27	Sequence 27, Appli
38	33	68.8	133	5	PCT-US93-11611-2	Sequence 2, Appli
39	33	68.8	133	5	PCT-US93-11611-9	Sequence 9, Appli
40	33	68.8	141	1	US-08-438-123-7	Sequence 7, Appli
41	33	68.8	726	4	US-09-302-812-10	Sequence 10, Appli
42	33	68.8	726	4	US-09-511-477-10	Sequence 10, Appli
43	33	68.8	726	4	US-09-511-507-10	Sequence 10, Appli
44	33	68.8	816	2	US-08-267-803B-9	Sequence 9, Appli
45	33	68.8	816	4	US-09-041-886-17	Sequence 17, Appli

## ALIGNMENTS

RESULT 1  
US-08-672-345C-27  
; Sequence 27, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-672-345C-27

Query Match 91.7%; Score 44; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.9e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QQLVEYPFT 9  
Db 1 QQPVEYPFT 9

RESULT 2

US-09-214-095D-27  
; Sequence 27, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 27  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-27

Query Match 91.7%; Score 44; DB 4; Length 113;  
Best Local Similarity 88.9%; Pred. No. 0.16; Mismatches 0; Indels 1; Gaps 0;  
Matches 8; Conservative 0;

Qy 1 QQLVEYPFT 9  
||| |||||  
Db 1 QQFVEYPFT 9

RESULT 3  
US-08-672-345C-8  
; Sequence 8, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-345C-8

Query Match 91.7%; Score 44; DB 2; Length 113;  
Best Local Similarity 88.9%; Pred. No. 0.16; Mismatches 0; Indels 1; Gaps 0;  
Matches 8; Conservative 0;

Qy 1 QQLVEYPFT 9  
||| |||||  
Db 94 QQFVEYPFT 102

RESULT 4  
US-08-672-345C-98  
; Sequence 98, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-98

Query Match 91.7%; Score 44; DB 2; Length 113;  
Best Local Similarity 88.9%; Pred. No. 0.16; Mismatches 0; Indels 1; Gaps 0;  
Matches 8; Conservative 0;

Qy 1 QQLVEYPFT 9  
||| |||||  
Db 94 QQFVEYPFT 102

RESULT 5  
US-09-214-095D-8  
; Sequence 8, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-8

Query Match 91.7%; Score 44; DB 4; Length 113;  
Best Local Similarity 88.9%; Pred. No. 0.16; Mismatches 0; Indels 1; Gaps 0;  
Matches 8; Conservative 0;

QY 1 QQLVEYPFT 9  
||| |||||  
Db 94 QQFVEYPFT 102

## RESULT 6

US-09-214-095D-108  
; Sequence 108, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 108  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Murine  
US-09-214-095D-108

Query Match 91.7%; Score 44; DB 4; Length 113;  
Best Local Similarity 88.9%; Pred. No. 0.16;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
||| |||||  
Db 94 QQFVEYPFT 102

## RESULT 7

US-08-672-345C-21  
; Sequence 21, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-21

Query Match 75.0%; Score 36; DB 2; Length 9;

Best Local Similarity 66.7%; Pred. No. 1.9e+05;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
||| |||||  
Db 1 QHFVDYPFT 9

## RESULT 8

US-09-214-095D-21  
; Sequence 21, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 21  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Murinae gen.sp.  
US-09-214-095D-21

Query Match 75.0%; Score 36; DB 4; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.9e+05;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPFT 9  
||| |||||  
Db 1 QHFVDYPFT 9

## RESULT 9

US-08-672-345C-5  
; Sequence 5, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

## US-08-672-345C-5

Query Match 75.0%; Score 36; DB 2; Length 113;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9  
Db 94 QHFVDYPT 102

## RESULT 10

US-08-672-345C-7  
; Sequence 7, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-345C-7

Query Match 75.0%; Score 36; DB 2; Length 113;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9  
Db 94 QHFVDYPT 102

## RESULT 11

US-08-672-345C-95  
; Sequence 95, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York

; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 95:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-345C-95

Query Match 75.0%; Score 36; DB 2; Length 113;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQLVEYPPT 9  
Db 94 QHFVDYPT 102

## RESULT 12

US-08-672-345C-97  
; Sequence 97, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 97:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: peptide  
US-08-672-345C-97

Query Match 75.0%; Score 36; DB 2; Length 113;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QQLVEYPFT 9  
| : |||||  
Db 94 QHFVDYPFT 102

## RESULT 13

US-09-214-095D-5  
; Sequence 5, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Murinae gen. Sp.  
US-09-214-095D-5

Query Match 75.0%; Score 36; DB 4; Length 113;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QQLVEYPFT 9  
| : |||||  
Db 94 QHFVDYPFT 102

## RESULT 14

US-09-214-095D-7  
; Sequence 7, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp  
US-09-214-095D-7

Query Match 75.0%; Score 36; DB 4; Length 113;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QQLVEYPFT 9  
| : |||||  
Db 94 QHFVDYPFT 102

## RESULT 15

US-09-214-095D-100  
; Sequence 100, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY

; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 100  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Murine  
US-09-214-095D-100

Query Match 75.0%; Score 36; DB 4; Length 113;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QQLVEYPFT 9  
| : |||||  
Db 94 QHFVDYPFT 102

Search completed: April 22, 2003, 12:55:41  
Job time : 7.875 secs





GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:49 ; Search time 14 seconds  
(without alignments)  
33.626 Million cell updates/sec

Title: US-09-674-716B-3

Perfect score: 81  
Sequence: 1 RSSKSLLYDKGKTYLN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/PCITUS COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77	95.1	16	2	US-08-672-345C-22
2	77	95.1	16	2	US-08-672-345C-25
3	77	95.1	16	4	US-09-214-095D-22
4	77	95.1	16	4	US-09-214-095D-25
5	77	95.1	113	2	US-08-672-345C-8
6	77	95.1	113	2	US-08-672-345C-98
7	77	95.1	113	4	US-09-214-095D-8
8	77	95.1	113	4	US-09-214-095D-108
9	75	92.6	16	2	US-08-672-345C-19
10	75	92.6	16	2	US-08-672-345C-28
11	75	92.6	16	4	US-09-214-095D-19
12	75	92.6	16	4	US-09-214-095D-28
13	75	92.6	113	2	US-08-672-345C-5
14	75	92.6	113	2	US-08-672-345C-6
15	75	92.6	113	2	US-08-672-345C-7
16	75	92.6	113	2	US-08-672-345C-95
17	75	92.6	113	2	US-08-672-345C-96
18	75	92.6	113	2	US-08-672-345C-97
19	75	92.6	113	4	US-09-214-095D-5
20	75	92.6	113	4	US-09-214-095D-6
21	75	92.6	113	4	US-09-214-095D-7
22	75	92.6	113	4	US-09-214-095D-100
23	75	92.6	113	4	US-09-214-095D-104
24	75	92.6	113	4	US-09-214-095D-112
25	75	92.6	280	4	US-09-214-095D-119
26	69	85.2	16	2	US-08-672-345C-43
27	69	85.2	16	2	US-08-672-345C-79

28	69	85.2	16	4	US-09-214-095D-43	Sequence 43, Appl
29	69	85.2	16	4	US-09-214-095D-79	Sequence 79, Appl
30	69	85.2	113	4	US-09-214-095D-116	Sequence 116, App
31	69	85.2	114	2	US-08-672-345C-9	Sequence 9, Appl
32	69	85.2	114	4	US-09-214-095D-9	Sequence 9, Appl
33	69	85.2	115	2	US-08-672-345C-99	Sequence 99, Appl
34	64	79.0	112	1	US-07-942-245-28	Sequence 28, Appl
35	59	72.8	112	1	US-08-477-877B-89	Sequence 89, Appl
36	59	72.8	112	2	US-08-472-281A-89	Sequence 89, Appl
37	59	72.8	112	2	US-08-678-194-6	Sequence 6, Appl
38	59	72.8	112	2	US-08-477-989B-89	Sequence 89, Appl
39	59	72.8	112	4	US-08-890-011-6	Sequence 6, Appl
40	59	72.8	112	4	US-09-262-724-6	Sequence 6, Appl
41	59	72.8	535	4	US-08-983-035A-38	Sequence 38, Appl
42	56	69.1	112	1	US-08-477-877B-87	Sequence 87, Appl
43	56	69.1	112	1	US-08-477-877B-88	Sequence 88, Appl
44	56	69.1	112	2	US-08-472-281A-87	Sequence 87, Appl
45	56	69.1	112	2	US-08-472-281A-88	Sequence 88, Appl

## ALIGNMENTS

RESULT 1  
US-08-672-345C-22  
; Sequence 22, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-22

Query Match 95.1%; Score 77; DB 2; Length 16;  
Best Local Similarity 93.8%; Pred. No. 1.3e-06;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RSSKSLLYDKGKTYLN 16

Db 1 RSSKSLLYDKGKTYLN 16

RESULT 2

```

US-08-672-345C-25
; Sequence 25, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-25

Query Match          95.1%; Score 77; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 1.3e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
DB 1 RSSKSLLYEDGKTYLN 16

RESULT 3
US-09-214-095D-22
; Sequence 22, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-22

Query Match          95.1%; Score 77; DB 4; Length 16;
Best Local Similarity 93.8%; Pred. No. 1.3e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
DB 1 RSSKSLLYEDGKTYLN 16

```

```

RESULT 4
US-09-214-095D-25
; Sequence 25, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-25

Query Match          95.1%; Score 77; DB 4; Length 16;
Best Local Similarity 93.8%; Pred. No. 1.3e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16
DB 1 RSSKSLLYEDGKTYLN 16

RESULT 5
US-08-672-345C-8
; Sequence 8, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-672-345C-8

Query Match          95.1%; Score 77; DB 2; Length 113;
Best Local Similarity 93.8%; Pred. No. 9.9e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 1 RSKSLLYKDGKTYLN 16
    |||||:|||||
Db 24 RSKSLLYEDGKTYLN 39

RESULT 6
US-08-672-345C-98
; Sequence 98, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-98

Query Match 95.1%; Score 77; DB 2; Length 113;
Best Local Similarity 93.8%; Pred. No. 9.9e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSKSLLYKDGKTYLN 16
    |||||:|||||
Db 24 RSKSLLYEDGKTYLN 39

RESULT 7
US-09-214-095D-8
; Sequence 8, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-8

Query Match 95.1%; Score 77; DB 4; Length 113;
Best Local Similarity 93.8%; Pred. No. 9.9e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSKSLLYKDGKTYLN 16
    |||||:|||||
Db 24 RSKSLLYEDGKTYLN 39

RESULT 8
US-09-214-095D-108
; Sequence 108, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murine
US-09-214-095D-108

Query Match 95.1%; Score 77; DB 4; Length 113;
Best Local Similarity 93.8%; Pred. No. 9.9e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSKSLLYKDGKTYLN 16
    |||||:|||||
Db 24 RSKSLLYEDGKTYLN 39

RESULT 9
US-08-672-345C-19
; Sequence 19, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-19

Query Match 95.1%; Score 77; DB 2; Length 113;
Best Local Similarity 93.8%; Pred. No. 9.9e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSKSLLYKDGKTYLN 16
    |||||:|||||
Db 24 RSKSLLYEDGKTYLN 39
```

US-08-672-345C-19

Query Match 92.6%; Score 75; DB 2; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.8e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSKSLLYKDGKTYLN 16  
Db 1 RSSRSLLYRDGKTYLN 16

RESULT 10

US-08-672-345C-28  
; Sequence 28, Application US/08672345C  
; Patent No. 5948658

GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper and Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-28

Query Match 92.6%; Score 75; DB 2; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.8e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSKSLLYKDGKTYLN 16  
Db 1 RSSRSLLYRDGKTYLN 16

RESULT 11

US-09-214-095D-19  
; Sequence 19, Application US/09214095D  
; Patent No. 6280987

GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214, 095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19

; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-19

Query Match 92.6%; Score 75; DB 4; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.8e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSKSLLYKDGKTYLN 16  
Db 1 RSSRSLLYRDGKTYLN 16

RESULT 12

US-09-214-095D-28  
; Sequence 28, Application US/09214095D  
; Patent No. 6280987

GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214, 095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 28

LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-28

Query Match 92.6%; Score 75; DB 4; Length 16;  
Best Local Similarity 87.5%; Pred. No. 2.8e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSKSLLYKDGKTYLN 16  
Db 1 RSSRSLLYRDGKTYLN 16

RESULT 13

US-08-672-345C-5  
; Sequence 5, Application US/08672345C  
; Patent No. 5948658

GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 5:

;  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-345C-5

Query Match 92.6%; Score 75; DB 2; Length 113;  
Best Local Similarity 87.5%; Pred. No. 2.1e-05;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
|||:||||:|||||  
Db 24 RSSRSLLYRDGKTYLN 39

## RESULT 14

US-08-672-345C-6  
; Sequence 6, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-345C-6

Query Match 92.6%; Score 75; DB 2; Length 113;  
Best Local Similarity 87.5%; Pred. No. 2.1e-05;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
|||:||||:|||||  
Db 24 RSSRSLLYRDGKTYLN 39

## RESULT 15

US-08-672-345C-7  
; Sequence 7, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY

;  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-345C-7

Query Match 92.6%; Score 75; DB 2; Length 113;  
Best Local Similarity 87.5%; Pred. No. 2.1e-05;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSSKSLLYKDGKTYLN 16  
|||:||||:|||||  
Db 24 RSSRSLLYRDGKTYLN 39

Search completed: April 22, 2003, 12:55:40  
Job time: 14 secs



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:48:49 ; Search time 6.125 seconds  
(without alignments)  
33.626 Million cell updates/sec

Title: US-09-674-716B-5  
Perfect score: 31  
Sequence: 1 LMSTRAS 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	7	2	US-08-672-345C-23
2	31	100.0	7	2	US-08-672-345C-26
3	31	100.0	7	2	US-08-672-345C-29
4	31	100.0	7	4	US-09-214-095D-23
5	31	100.0	7	4	US-09-214-095D-26
6	31	100.0	7	4	US-09-214-095D-29
7	31	100.0	113	2	US-08-672-345C-6
8	31	100.0	113	2	US-08-672-345C-7
9	31	100.0	113	2	US-08-672-345C-8
10	31	100.0	113	2	US-08-672-345C-96
11	31	100.0	113	2	US-08-672-345C-97
12	31	100.0	113	2	US-08-672-345C-98
13	31	100.0	113	4	US-09-214-095D-6
14	31	100.0	113	4	US-09-214-095D-7
15	31	100.0	113	4	US-09-214-095D-8
16	31	100.0	113	4	US-09-214-095D-104
17	31	100.0	113	4	US-09-214-095D-108
18	31	100.0	113	4	US-09-214-095D-112
19	28	90.3	7	2	US-08-672-345C-20
20	28	90.3	7	4	US-09-214-095D-20
21	28	90.3	113	2	US-08-672-345C-5
22	28	90.3	113	2	US-08-672-345C-95
23	28	90.3	113	4	US-09-214-095D-5
24	28	90.3	113	4	US-09-214-095D-100
25	28	90.3	280	4	US-09-214-095D-119
26	27	87.1	7	2	US-08-672-345C-80
27	27	87.1	7	4	US-09-214-095D-80

28	26	83.9	250	4	US-09-029-348-14	Sequence 14, Appl
29	26	83.9	1341	3	US-08-963-825-18	Sequence 18, Appl
30	26	83.9	1341	4	US-09-500-811-18	Sequence 18, Appl
31	26	83.9	1341	4	US-09-570-573-18	Sequence 18, Appl
32	26	83.9	1341	4	US-09-548-608-18	Sequence 18, Appl
33	26	83.9	1461	4	US-09-585-887-9	Sequence 9, Appl
34	26	83.9	1461	4	US-09-289-578-9	Sequence 9, Appl
35	25	80.6	664	6	5240838-5	Patent No. 5240838
36	24	77.4	274	4	US-09-188-930-336	Sequence 336, Appl
37	24	77.4	443	2	US-08-833-963C-2	Sequence 2, Appl
38	24	77.4	443	3	US-08-980-514-1	Sequence 1, Appl
39	24	77.4	534	4	US-09-029-348-5	Sequence 5, Appl
40	24	77.4	537	4	US-09-029-348-4	Sequence 4, Appl
41	24	77.4	626	4	US-09-029-348-2	Sequence 2, Appl
42	24	77.4	801	4	US-09-104-070-2	Sequence 2, Appl
43	24	77.4	1481	2	US-08-616-844-40	Sequence 40, Appl
44	24	77.4	1481	2	US-08-599-654-40	Sequence 40, Appl
45	24	77.4	1481	3	US-08-944-868A-40	Sequence 40, Appl

## ALIGNMENTS

RESULT 1  
US-08-672-345C-23  
; Sequence 23, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-672-345C-23

Query Match 100.0%; Score 31; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7  
| | | | | | |  
Db 1 LMSTRAS 7

RESULT 2

```
US-08-672-345C-26
; Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-26

Query Match 100.0%; Score 31; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 1 LMSTRAS 7

RESULT 3
US-08-672-345C-29
; Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-26

Query Match 100.0%; Score 31; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 1 LMSTRAS 7

RESULT 3
US-08-672-345C-29
; Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-29

Query Match 100.0%; Score 31; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 1 LMSTRAS 7

RESULT 5
US-09-214-095D-26
; Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-23

Query Match 100.0%; Score 31; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 1 LMSTRAS 7

RESULT 5
US-09-214-095D-26
; Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-26

Query Match 100.0%; Score 31; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



QY 1 LMSTRAS 7  
| | | | |  
Db 1 LMSTRAS 7

## RESULT 6

US-09-214-095D-29  
; Sequence 29, Application US/09214095D  
; Patent No. 6280987  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; FILE REFERENCE: 51400-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/214,095D  
; CURRENT FILING DATE: 1999-07-19  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 29  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Murinae gen. sp.  
US-09-214-095D-29

Query Match 100.0%; Score 31; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
| | | | |  
Db 1 LMSTRAS 7

## RESULT 7

US-08-672-345C-6  
; Sequence 6, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-345C-6

Query Match 100.0%; Score 31; DB 2; Length 113;

Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
| | | | |  
Db 55 LMSTRAS 61

## RESULT 8

US-08-672-345C-7  
; Sequence 7, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,345C  
; FILING DATE: 24-JUN-1996  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-345C-7

Query Match 100.0%; Score 31; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
| | | | |  
Db 55 LMSTRAS 61

## RESULT 9

US-08-672-345C-8  
; Sequence 8, Application US/08672345C  
; Patent No. 5948658  
; GENERAL INFORMATION:  
; APPLICANT: Landry, Donald, W.  
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper and Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patent In Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/672,345C  
/ FILING DATE: 24-JUN-1996  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: White, John P.  
/ REGISTRATION NUMBER: 28,678  
/ REFERENCE/DOCKET NUMBER: 0575/51400  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 212-278-0400  
/ TELEFAX: 212-391-0525  
/ INFORMATION FOR SEQ ID NO: 8:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 113 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ US-08-672-345C-8

Query Match 100.0%; Score 31; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
Db 55 LMSTRAS 61

RESULT 10  
US-08-672-345C-96  
/ Sequence 96, Application US/08672345C  
/ Patent No. 5948658  
/ GENERAL INFORMATION:  
/ APPLICANT: Landry Donald, W.  
/ TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
/ NUMBER OF SEQUENCES: 108  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Cooper and Dunham LLP  
/ STREET: 1185 Avenue of the Americas  
/ CITY: New York  
/ STATE: New York  
/ COUNTRY: USA  
/ ZIP: 10036  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patent In Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/672,345C  
/ FILING DATE: 24-JUN-1996  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: White, John P.  
/ REGISTRATION NUMBER: 28,678  
/ REFERENCE/DOCKET NUMBER: 0575/51400  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 212-278-0400  
/ TELEFAX: 212-391-0525  
/ INFORMATION FOR SEQ ID NO: 96:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 113 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ US-08-672-345C-96

Query Match 100.0%; Score 31; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
Db 55 LMSTRAS 61

RESULT 11  
US-08-672-345C-97  
/ Sequence 97, Application US/08672345C  
/ Patent No. 5948658  
/ GENERAL INFORMATION:  
/ APPLICANT: Landry Donald, W.  
/ TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
/ NUMBER OF SEQUENCES: 108  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Cooper and Dunham LLP  
/ STREET: 1185 Avenue of the Americas  
/ CITY: New York  
/ STATE: New York  
/ COUNTRY: USA  
/ ZIP: 10036  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patent In Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/672,345C  
/ FILING DATE: 24-JUN-1996  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: White, John P.  
/ REGISTRATION NUMBER: 28,678  
/ REFERENCE/DOCKET NUMBER: 0575/51400  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 212-278-0400  
/ TELEFAX: 212-391-0525  
/ INFORMATION FOR SEQ ID NO: 97:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 113 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ US-08-672-345C-97

Query Match 100.0%; Score 31; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMSTRAS 7  
Db 55 LMSTRAS 61

RESULT 12  
US-08-672-345C-98  
/ Sequence 98, Application US/08672345C  
/ Patent No. 5948658  
/ GENERAL INFORMATION:  
/ APPLICANT: Landry Donald, W.  
/ TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY  
/ NUMBER OF SEQUENCES: 108  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Cooper and Dunham LLP  
/ STREET: 1185 Avenue of the Americas  
/ CITY: New York  
/ STATE: New York  
/ COUNTRY: USA  
/ ZIP: 10036

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-672-345C-98

Query Match 100.0%; Score 31; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 55 LMSTRAS 61

RESULT 13
US-09-214-095D-6
; Sequence 6, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; US-09-214-095D-6

Query Match 100.0%; Score 31; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 55 LMSTRAS 61

RESULT 14
US-09-214-095D-7
; Sequence 7, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; US-09-214-095D-7

Query Match 100.0%; Score 31; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 55 LMSTRAS 61

RESULT 15
US-09-214-095D-8
; Sequence 8, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; US-09-214-095D-8

Query Match 100.0%; Score 31; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMSTRAS 7
Db 55 LMSTRAS 61

Search completed: April 22, 2003, 12:55:41
Job time : 7.125 secs
```

